


```

**** 7 standard deviations above mean ****
3. 047834 Murine interleukin 9 receptor 2281 134 667 7.52 0
4. 003684 PAT144 clone of gene encoding 659 133 251 7.46 0
5. 003748 EPO receptor gene. 1883 133 653 7.46 0
6. 023129 Sequence encodes human MIP-1b 352 132 241 7.39 0
7. 003297 Act-2 clone of gene encoding 711 132 261 7.39 0
8. 003682 Act-2 clone of gene encoding 711 132 259 7.39 0
9. 028806 DNA encoding placental protei 820 132 368 7.39 0
10. 029260 Human calcium channel 27980/2 4919 129 752 7.19 0
11. 006470 Sequence encoding complete fo 3086 128 756 7.13 0
**** 6 standard deviations above mean ****
12. 024676 Encodes partial MP1V-env rela 552 124 265 6.86 0
13. 057445 Phenylethanolamine-N-Me-trans 792 124 352 6.86 0
14. 024675 Encodes MP1V-env related poly 1307 120 381 6.60 0
15. 040441 Sequence of a cDNA clone cont 6314 120 751 6.60 0
16. 024322 Mutant thermostable DNA polym 1635 117 697 6.40 0
17. N30043 Sequence encoding bovine leuk 1182 114 509 6.20 0
18. 013332 GDP-Fuc:beta-D-galactoside al 8174 114 763 6.14 0
19. N60292 Intron free glucosylase stru 2016 113 769 6.14 0
20. N60229 Mouse tumour necrosis factor 3500 113 782 6.14 0
21. N30044 Sequence encoding bovine leuk 1182 112 486 6.07 0
22. 004781 Sequence encoding Ty3 positio 1350 112 281 6.07 0
23. 024321 Mutant thermostable DNA polym 1899 112 717 6.07 0
24. 024320 Mutant thermostable DNA polym 2043 112 717 6.07 0
25. 024013 Mutant thermostable DNA polym 2277 112 717 6.07 0
26. 024012 Mutant thermostable DNA polym 2370 112 717 6.07 0
27. 024011 Mutant thermostable DNA polym 2505 112 717 6.07 0
28. N60294 C-terminal of glucosylase ge 1753 111 495 6.01 0
29. 022696 Gal alpha-2,6-ST (clone 14). 2188 111 625 6.01 0
**** 5 standard deviations above mean ****
30. N90501 Goat growth hormone precursor 817 110 315 5.94 0
31. N60274 Sequence encoding creatine am 1212 110 534 5.94 0
32. N81271 Sequence encoding creatine am 1212 110 532 5.94 0
33. 036508 DNA encoding a glycosyltransf 8174 110 763 5.94 0
34. 024677 Encodes MP1V env-vmp1 fusion 1216 109 452 5.87 0
35. 043995 IDUA - exons III to XIV. 4475 109 767 5.87 0
36. 040862 VAC-beta (genomic - part A). 13633 109 758 5.87 0
37. 003111 DNA sequence of gene encoding 1559 108 205 5.81 0
38. 006689 5' UTR of ELAM1 from clone EL 1566 108 262 5.81 0
39. 020727 Proteinase 3 coding sequence. 799 107 360 5.74 0
40. 013616 VNTR locus D20S15 flanking re 1284 107 323 5.74 0
41. N80113 Encodes recombinant human lys 1661 107 591 5.74 0
42. 004490 Sequence encoding human bone 1685 107 723 5.74 0
43. 046161 DNA encoding tumour associate 2206 107 640 5.74 0
44. 046888 Sequence encoding an immune 2206 107 640 5.74 0
45. 048170 Sequence encoding a 90K tumou 2206 107 639 5.74 0

```

1. US-08-223-263-2 (1-1795)
000359 Pseudorabies virus glycoprotein gII polypeptide.

ID 003259 standard; DNA; 4897 BP.
AC 003259;
DT 22-JUL-1990 (first entry)
DE Pseudorabies virus glycoprotein gII polypeptide.
KW Aujeszky's disease; vaccine; herpes virus; pig; ss.
OS Pseudorabies virus.
FT Key Location/Qualifiers
CDS 2639..3001
FT /tag= a

```

FT /note="gII glycoprotein"
PN EP-353809-A.
PD 07-FEB-1990.
PF 18-JUL-1989; 201887.
PR 01-AUG-1988; EP-112479.
PA (AIKU) Akzo NV.
PI Schreurs CS, Mettemleiter TC, Simon AJ, Lukacs N, Rhiza HJ;
DR MPI; 90-038534/06.
DR P-PSDB; R03732.
PT Immunogenic polypeptide of pseudo-rabies virus -
PT for use in mammalian vaccine against Aujeszky's disease.
PS Disclosure; Fig 1; 21pp; English.
CC The polypeptide may be used as a vaccine to provide
CC mammalian, e.g. pigs, with resistance against Aujeszky's
CC disease.
SQ Sequence 4897 BP; 594 A; 2052 C; 1495 G; 756 T;

Initial Score = 162 Optimized Score = 730 Significance = 9.37
Residue Identity = 46% Matches = 867 Mismatches = 778
Gaps = 202 Conservative Substitutions = 0

X 10 20 30 40 50
TCCTTCATCCATCTGCTCCACAGG--GCTGCTGC--TGTGCACTTGG--TCTTGA
ATGCCACTTTTTCACAGGCGCTGCGACGCTGGCGCGCGCGCTGCGCAAGTGTCTGGCGCGCGG
3150 X 3160 3170 3180 3190 3200 3210

60 70 80 90 100 110
GCCCT--TCTCCAGC--GGG--ATAGATTCCTACACCTTGGCC--CGCTTGGCCCAACCTAC--TCTGC
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
GCCGATCTCGCGCGCTGGGGGAGATGTCCT--TCT--GTCCAAACCTTGGGGGCTGCGCATCGGG
3220 3230 3240 3250 3260 3270 3280

120 130 140 150 160 170 180
CCGAAGTCCAG--AGCT--AAAGCCCTCATGCGC--CCAAGGAAGATTCAGGGGAGAGCCCAAGCA
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
CTGCTGCTCGCGCGCGCTGGCTGCGGCTTCTGCGCTTACAGGCAATCTGCGCTGCGCTGCG--AAC-
3290 3300 3310 3320 3330 3340 3350

190 200 210 220 230 240 250
GGAGGCGACGCGACGACGACCCCGGCGCAGA--ATGAGGCTGATGATGCTCTGCTGCTCATGCTTCT
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
----CCATGAAGGCGCTGATCCCGCTCAGCAGCAAGCGCTAAGSA---GGAGGCGCTCAGCAG--C
3360 3370 3380 3390 3400 3410

260 270 280 290 300 310
CCTAATCG--CAAGGCTAAGCT--GTCCA--GCCCGG--CTCTCTGCTGCTTGTGACCTCCAGTCTCATGA
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
GCGACGTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG
3420 3430 3440 3450 3460 3470 3480

320 330 340 350 360 370 380
AACTGCTTGTGATCCATCCATCTGCTTACAGCAG--ACTGAGCCAGTCCAGAGTTCACCTTTGCCATCA
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
GA--GGAGCAGGA--GCACAGAGCGCGCA--AGAAAGACAGCGCGCGCGCTGCTGCGCAGCGCGTCCGGCG
3490 3500 3510 3520 3530 3540 3550

390 400 410 420 430 440 450
CCTGTC--TCTGCTGCTGCTGCTGCTGCTTGTGAGAGAGTGAAGAACCCAGATG--GAGAGACCAAGG
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111

```

[illegible][illegible]

2. US-08-223-263-2 (1-1795)
Q27816 Adrenalin receptor subtype beta-1 gene.

Accession	Gene	Chromosome	Start (kb)	End (kb)	Orientation	Strand	Length (bp)	GC Content (%)	GC Skew	GC Bias	GC Bias2	GC Bias3	GC Bias4	GC Bias5	GC Bias6	GC Bias7	GC Bias8	GC Bias9	GC Bias10	GC Bias11	GC Bias12	GC Bias13	GC Bias14	GC Bias15	GC Bias16	GC Bias17	GC Bias18	GC Bias19	GC Bias20	GC Bias21	GC Bias22	GC Bias23	GC Bias24	GC Bias25	GC Bias26	GC Bias27	GC Bias28	GC Bias29	GC Bias30	GC Bias31	GC Bias32	GC Bias33	GC Bias34	GC Bias35	GC Bias36	GC Bias37	GC Bias38	GC Bias39	GC Bias40	GC Bias41	GC Bias42	GC Bias43	GC Bias44	GC Bias45	GC Bias46	GC Bias47	GC Bias48	GC Bias49	GC Bias50	GC Bias51	GC Bias52	GC Bias53	GC Bias54	GC Bias55	GC Bias56	GC Bias57	GC Bias58	GC Bias59	GC Bias60	GC Bias61	GC Bias62	GC Bias63	GC Bias64	GC Bias65	GC Bias66	GC Bias67	GC Bias68	GC Bias69	GC Bias70	GC Bias71	GC Bias72	GC Bias73	GC Bias74	GC Bias75	GC Bias76	GC Bias77	GC Bias78	GC Bias79	GC Bias80	GC Bias81	GC Bias82	GC Bias83	GC Bias84	GC Bias85	GC Bias86	GC Bias87	GC Bias88	GC Bias89	GC Bias90	GC Bias91	GC Bias92	GC Bias93	GC Bias94	GC Bias95	GC Bias96	GC Bias97	GC Bias98	GC Bias99	GC Bias100
1100	1110	1120	1130	1140	1150	1160	1170	1180	1190	1200	1210	1220	1230	1240	1250	1260	1270	1280	1290	1300	1310	1320	1330	1340	1350	1360	1370	1380	1390	1400	1410	1420	1430	1440	1450	1460	1470	1480	1490	1500	1510	1520	1530	1540	1550	1560	1570	1580	1590	1600	1610	1620	1630	1640	1650	1660	1670	1680	1690	1700	1710	1720	1730	1740	1750	1760	1770	1780	1790	1800	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000																			
ACAGACA-CTGGGACATCATGATGATGGTGTTCACGCT-CTTCTGGTGGCCCTTCTTCTGAGCA--ACGTA	940	950	960	970	980	990	1000	1010	1020	1030	1040	1050	1060	1070	1080	1090	1100	1110	1120	1130	1140	1150	1160	1170	1180	1190	1200	1210	1220	1230	1240	1250	1260	1270	1280	1290	1300	1310	1320	1330	1340	1350	1360	1370	1380	1390	1400	1410	1420	1430	1440	1450	1460	1470	1480	1490	1500	1510	1520	1530	1540	1550	1560	1570	1580	1590	1600	1610	1620	1630	1640	1650	1660	1670	1680	1690	1700	1710	1720	1730	1740	1750	1760	1770	1780	1790	1800	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000		
CCCCAACCATCTCTC--CTAC-TGG-ACAGTATAGCTCTTCCCTCTTC--CAACCACCTTCCGAC-CGCT	1100	1110	1120	1130	1140	1150	1160	1170	1180	1190	1200	1210	1220	1230	1240	1250	1260	1270	1280	1290	1300	1310	1320	1330	1340	1350	1360	1370	1380	1390	1400	1410	1420	1430	1440	1450	1460	1470	1480	1490	1500	1510	1520	1530	1540	1550	1560	1570	1580	1590	1600	1610	1620	1630	1640	1650	1660	1670	1680	1690	1700	1710	1720	1730	1740	1750	1760	1770	1780	1790	1800	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000																		
GTGAAGCTTTCACGCGACCTGGTGGCGGATGCGCTT-CGCTTTCTTCAAGTGGCTGGAGTACGCAAA	1010	1020	1030	1040	1050	1060	1070	1080	1090	1100	1110	1120	1130	1140	1150	1160	1170	1180	1190	1200	1210	1220	1230	1240	1250	1260	1270	1280	1290	1300	1310	1320	1330	1340	1350	1360	1370	1380	1390	1400	1																																																																				

[illegible]

[illegible][illegible]

Residue Identity =	50%	Matches =	798	Mismatches =	564
Gaps =	208	Conservative Substitutions =			0

TCT-TCTTTC-CCATCTCTCTCCCAAGGGCT--GGCTGCTGTGACTTT-GGGTCTGTGAGCT
 ---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
 ATGGAAGCTGTGTCGTGCTACCA---GGCTCCC---ACGGCTGTGATGATGGGTGCGTTGTGTGTCCTGAT
 410 X 420 430 440 450 460 470

60 70 80 90 100 110 120
CCCTC--TCCACCCGGATGATTCCTACCCCTTGGCCCGC-CTTTCGCCCAACCTACTCTGCCCAATGCG
CCTACGCCACCGACCTTCGACCTTCGCCCTTACGATTGGGCGCTACACGACCGCTCCGGGCGCCGATAT

130 140 150 160 170 180 190
AAGAGCCGTAAGCCGCCCTCCATCGAGCCCGAGGATTCAGGAGGAGGAGGCCCAACAG--GAGACCAACCCCA
CAGCGTGTATCCATCATCATG----AGTGTGCTCTCTAG--AGGCCCCCGGTGGGAGCTGTGTGGCGCGGTTT
550 560 570 580 590 600

200 210 220 230 240 250 260
 GCACAC-ACCCGGCCAGATGAGCG-ACTAATTCCTCTCTGTC-GTC----ATG-CTTCTCCACAC
 |||||
 610 620 630 640 650 660 670
 GCGTTCAGAGAGCGGCCA-CGTATGTTTCCTGCTGCTCCGCGCCCTGAGACCAACCATGCGTTCATCCTCC

[illegible][illegible]

400
GCT--GCC-TGCT--GTGGACTTATGCTGGAGAAATGG-AAAACCAATG-----GAGAGAACAAG-GC
820
GCTATGCCAGACTTGTCGCGCGCT---GTGGAGCCCTGTGTGGAGCTTG-TGTGCCTCTGAAGCCCTACCA

460 470 480 490 500 510 520
ACAGGACTTCTGGAGCAGTAGCCCTTCCTGGAGGATGATGGACGACCGGGCAACCTGGACC-CA
||||| ||||| ||||| ||||| ||||| |||||
CTGGACCCCCT--CATC-CTGACGCTTCCCTCATCTGTG--GTATTC-CTGGTGCTGC--GACCGTG

530 540 550 560 570 580 590
CTTGG-CT-CTCAATCCTCTCTCTGGGACGCTTTTGGG-CA-----GGTCTCTCCTTCTGGGGCCCTG
||||| ||||| ||||| ||||| ||||| ||||| |||||
CTGGGCGCTGCTC-TCGACCGCGGG-GG-TCTGAAGCAGAAATCTCTGCTGCTGCTGATCC-CAACACAG-
960 970 980 990 1000 1010

AGAGCCCTCCTGACACCAGCTT--CTTC-AC-AGGGACAGACACACT-CACAAGATCCCAATGCATT
||||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||

AGAG-CGAGTTGAAGSCTCTTCACCACCCAAGGGTACTTC-CAGCTGTGGCTGTACCAAGATATGGTGGT

660 670 680 690 700 710 720 730
CTTCCAGGCTTCACAAACACTCTCCAGGAAAGTGCGTTTCTGATGGCTGTGAAGAGGATCAACCCTC
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
CTGCCTGTG-GTGAAGCCCCTCACACCCCCTTCACG-GAAGACCACTGCTCCCTGGAGTCTCTCAGAC

740 750 760 770 780 790
 CS--TAAAGGAG--GCCCAACCCACACAGCTGCCACAGCA----ACCTCTCTCTCTCCTACACT-----
 1160 1170 1180 1190 1200 1210 1220
 CCTCTCTGGGGGACATTCACAGCAGCTGGAGCGCGGGGACACATATATAGAGGCCCTCTCTGTGACCGCATGGGC

AACGAGC 800 810 820 830 840 850
 AACGAGC--TCCCAACGAGACT--TCGTGATGGTGGAGAAATCTACGCCTC--AGCGAACTACTG-
 1230 1240 1250 1260 1270 1280 1290
 AGTAGCATGCCAGGATACCTATCTGG--TCTCGACA--ATAGTTCCTGCCCGGAAACCC--CCCATGTG-

[illegible]

9350 940 950 960 970 980 990
CGAGTC-CCTGGACCAATCCCGGAGAACCTGAACGATACCAATCTTGAATGAAAC-TGCTGACCTC
CCTGCTCATCTG-CCTTGGCCTTCACAGCCAGCCCA-GAGGAA-CC-CTC-TGCTGCCACCTTATGATACAC
1360 1370 1380 1390 1400 1410 1420

1000 1010 1020 1030 1040 1050 1060
 CTTTCTGGAGCC--TCAGAGGAGCCATGAGACCCGGGACG--TTTCTCGAAGACATGACAGAGCTT
 1400 1410 1420 1430 1440 1450 1460 1470 1480
 CTATCTGAGCCAGCCTCCAG--CTTTGCGTCATGAGACCTTGCCCTG--AGTGGCCC--CTT

1070	1080	1090	1100	1110	1120	1130
CGTGCACCCACCTCTCAG--CCTGGATTTCTCTTCCCAACCATCTTC--CTACGTGCAGTATACGCT						
CC--CCACCCACCTTAAAGACTCTACCTGTGTGATCTGTGCATCTGGCATCTCACT--TACA--GCT						
1490	1500	1510	1520	1530	1540	1550

[illegible][illegible]

1270 1280 1290 1300 1310 1320
 CTCACGAAGGCTAAGTCTCAGA----CAGTCCGACATCAGCAT--TGTCTCATGTACAGC--TCCT-T
 ||||| | ||||| | ||||| | | | | | | | | |

KW	haematopoietic cell; Growth Factor receptor; ss
OS	Myeloproliferative leukaemia virus.
FH	Key
FT	Location/Qualifiers
CDS	1..1149

FN /*tag= a
PT WO9207074-A.
PD 30-APR-1992.
PF 19-OCT-1990. F00762.
PR 19-OCT-1990. WO-F07623.
PI (INRM) INSERM INT NAT SANTE & RECH MEDICALE.
PA Chabron M, Giselbrecht S, Penotelli JF, Souyri M,
PI Tambourin P, Varlet P, Vigon I, Wendling F,
DR WEF, 92-167154/20.

PT Polypeptides similar to v-mlp protein of MPLV - for diagnosis and
PT treatment of myeloproliferative diseases
Claim 10; Page 52-56; 75pp; French.
PS

CC This nucleotide sequence is one of 6 claimed sequences which code
CC for polypeptides related to MPV1 env protein. The proteins have one
CC or more of the following properties: they are involved in the
CC ligand-fixing or signal-transmitting function of haematopoietic
CC growth factor receptors; they are recognised by antibodies to the
CC protein with GENESEQ accession number R23970; when produced from
CC the MPV1 genome they can induce/promote proliferation of
CC haematopoietic cells and/or are involved in differentiation of
CC haematopoietic cells. The precise identity of sequence Q2675 is
CC not further described in the specification.

Sequence 1307 BP; 269 A; 435 C; 327 G; 276 T;

Initial Score	=	120	Optimized Score	=	381	Significance	=	6.60
Residue Identity	=	53%	Matches	=	471	Mismatches	=	269
Gaps	=	134	Conservative Substitutions				=	0

X 10 20 30 40 50
 TCTTCTACG-CATCTGCTCCAGAGGGGTGCCTCTCTGTGAATTG--GTTCTTGAACCCC
 | | | | |
 ACCTCAGAAGCTCTGCTACCAAGTCGGTAACCGGGAAGGCC--GTGAGACTGGAAGSTCTTGACCAC
 500 X 510 520 530 540 550 560

[illegible][illegible]

	200	210	220	230	240	250	260
C-GCCAGCCGACGCCCG-CGAGATGGACTGAATTCTCTCGTGCATCATTCTTA							
CGGCTCCGAGACTGTTCGATCACTTG---	TGACTGCTGTCGTCGATGCTGA- GCTCAGTGCCTT						
700	710	720	730	740	750	760	

270 280 290 300 310 320
CAAG-CTAAGCCTTCACAGCCCGGC-TTCCTCGCTTGACCTCC--GAGTCCTCAGTAAACTGCTTCGT

CTGGGCTACTGCTG-CTAAGTGGCAATTTCCTGC-----GCACTACAGAG-ACGTAGGCA--TGCCTTCT
770 780 790 800 810 820

330 340 350 360 370 380 390
G-ACTCCATC---GTTCCTCAGACAGACTGACGACGGC-CAGAGGTCACACCTTTCCTACACTC-CAGAGGTCACACCTTTCCTACACTC
|||||
GGCCTCGCTTCACAGACTTCACCGGGCTCTTGCGCCATGACTCTCAGAC-ACAC---TGC-AGCCCTMAAG
830 840 850 860 870 880 890

400 410 420 430 440 450 460
TCTCTCT--GCTCTCTTGACATTTAGCTTGGAGAA--TGGAAACCCAG--ATGGAGAACCAAGACACAA
TTTTTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT
TCTCTTAGGCACAGGTACCGATTACGTGTGACAGAGTGG--AACCCAGGCTCTTGGAAAAATTCCTCCCA

[illegible][illegible][illegible][illegible]

730	740	750	760	770	780	790
CTGGCTCAGGGGGGCCCCAC	-CCACAC-	AGCTGTCCACGAGAAC	-CTCTTAG-	-----	TCTC	CACT
GT-TGCCA-----	CCCCACTCCCCCTGGCTTTT	TAACTGATCACTC	-CAGATGCTGCTCAAAA			
1230	1240	1250	1260	1270	1280	

800 810 820
AAGAGCTCCAAACAGAGCTTGTGATTG
||||| | |
CCAGAGCTGTGCTCTGACG
1290 1300 X

15.	US-08-223-263-2 (1-1795)	
Q40341	Sequence of a cDNA clone contg. the ADMLX gene	ass
ID	Q40341 standard; cDNA to mRNA; 6314 BP.	
NC	Q40341.	

DT	09-AUG-1993	(first entry)
DE	Sequence of a cDNA clone contg. the ADMLX gene associated with X-	

DE linked Kallmann Syndrome (KS) .
 KM X-linked Kallmann syndrome; ADMX gene; diagnosis; PCR; ss.
 OS Homo sapiens. Location/Qualifiers
 FH CDS
 FT key
 FT /tag= a
 FT exon
 FT /tag= b
 FT /label= 1
 FT exon
 FT /tag= c
 FT /label= 2
 FT exon
 FT /tag= d
 FT /label= 3
 FT exon
 FT /tag= e
 FT /label= 4
 FT exon
 FT /tag= f
 FT /label= 5
 FT exon
 FT /tag= g
 FT /label= 6
 FT exon
 FT /tag= h
 FT /label= 7
 FT exon
 FT /tag= i
 FT /label= 8
 FT exon
 FT /tag= j
 FT /label= 9
 FT exon
 FT /tag= k
 FT /label= 10
 FT exon
 FT /tag= l
 FT /label= 11
 FT exon
 FT /tag= m
 FT /label= 12
 FT exon
 FT /tag= n
 FT /label= 13
 FT exon
 FT /tag= o
 FT /label= 14
 PN MO307267-A.
 PD 15-APR-1993.
 PF 09-OCT-1992; F00956.
 PR 09-OCT-1991; FR-012451.
 PA (INSP) INST PASTEUR.
 PI (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Claverie J, Hardelin J, Legouis R, Leuilliers J,
 PI Intalla G, Petit C,
 PI WPI93-134456/16.
 DR P-PDB; R34445.
 PT Nucleic acid sequence of gene with X-linked Kallmann syndrome -
 useful for diagnosing Kallmann syndrome by amplification to

PT detect genetic anomalies
 PS Claim 1; Fig 6; 60pp; French.
 CC The nucleic acid sequence is derived from the ADMX gene associated
 CC with KS (or hypogonadotropic hypogonadism and anosmia).
 CC Oligonucleotide pairs which act as primers for specific amplification
 CC of the gene are used in amplification methods to detect genetic
 CC anomalies which cause KS. The primer pairs corresp. to the coding
 CC and non-coding regions of exon 1 of the ADMX gene and one pair each
 CC for the other 13 exons.
 SQ Sequence 6314 Bp; 1884 A; 1274 C; 1246 G; 1910 T;
 Initial Score = 120 Optimized Score = 751 Significance = 6.60
 Residue Identity = 47% Matches = 922 Mismatches = 748
 Gaps = 251 Conservative Substitutions = 0
 X
 120 10 20 30 40 50
 CTCTCCACCCGAGATGATTTCTTCCACCCCTTGCCCGCCCTTGGCCACCCCTACTCTG-CCCAAGATGACAG
 60
 1270 1280 1290 1300 1310 1320 1330
 ATTT-GCAAGCC-ATAACGTAAGTGGGACACAGACGCG-AGAAGTGCAGAAAGTGTCTCTTCACTT-AC
 CTACGATGGGTTTCAGAAATTTCTGATCTGAGAGAACTCCAGCCAGACTGTG-AC-TATGTTTGGA-
 1210 X 1220 1230 1240 1250 1260
 TCTTCCATCCATCTGCTC-CCAG-AGGGCTGCTG-CTGTCACTTGGTCTGAGCC
 60
 70 80 90 100 110 120
 CTCTCCACCCGAGATGATTTCTTCCACCCCTTGCCCGCCCTTGGCCACCCCTACTCTG-CCCAAGATGACAG
 1270 1280 1290 1300 1310 1320 1330
 ATTT-GCAAGCC-ATAACGTAAGTGGGACACAGACGCG-AGAAGTGCAGAAAGTGTCTCTTCACTT-AC
 1340 1350 1360 1370 1380 1390 1400
 ATGACATGATGACACCAACAAAGAACAGCTT-GTGAA-ACAGAAAGTGAATTCAAACACA
 1410 1420 1430 1440 1450 1460 1470
 ACACCCCGGCGAATGAGTGAATGCTC-CTGTGTC-ATGCT-TCTCTTCACTGCAAGCTTAA
 210 220 230 240 250 260 270
 ACTCCCTTTTC-AAAGACGACGCCCTGCGCGTGAATGCGAGCTCTTCTTACAGATGGCCAA
 1410 1420 1430 1440 1450 1460 1470
 CGCTGTCCAGCCCGCTCTCTCTGCTTG-TAACCTCCAGTCTCACTGTAATGCT-TCTGTATCCCATG-
 280 290 300 310 320 330
 --CTG-CAAGTTAAAGT-CTACTGCAAGAGAGCAAGATCCCACTGTCAACGATATCATG--TCCGCTG
 1480 1490 1500 1510 1520 1530
 340 350 360 370 380 390
 --TCTTTCACAG-----CA-GACTGAGCCA-GTGCCAGAGG-TTGACC--CTTGTCCATC-ACCT--
 1540 1550 1560 1570 1580 1590 1600
 TTTCCTGAAGCGGTGCCCCACACAGAACAAACCGGATGACGATCATCTGCGATCACAGAAATTTAC
 400 410 420 430 440 450
 GTCTCTGCTGCTCTGTGACTTGAAGTTGTTGGAGAGATGAAACCCAGATGAGAGACCAA--GGC--AC
 1610 1620 1630 1640 1650 1660 1670
 ATATATCTTCAGATCT-CTCATTTCTTCAGCAAGAT--AA--GGTACCTGTCA-ACCAATATCCGCAAA
 460 470 480 490 500 510 520
 AGACATTC--TGGAGCAG-TG-ACCCTT--CTCTGTGAGGAGTATGACAGACGGGAGCACTGTGAC
 1610 1620 1630 1640 1650 1660 1670

FastDB - Fast Pairwise Comparison of Sequences

Release 5.4

Results file sq2.res made by on Tue 11 Apr 95 10:59:11-PDT.

```
Query sequence being compared: US-08-223-263-2 (1-1795)
Number of sequences searched: 358124
Number of scores above cutoff: 4841
```

Results of the initial comparison of US-08-223-263-2 (1-1795) with

Category	Value
1000000	1000000
500000	500000
100000	100000
50000	50000
10000	10000
5000	5000
1000	1000
500	500
100	100
50	50

Listing for Mary Hale

Tue Apr 11 13:28:23 1995

Page
42

[illegible]

Similarity matrix	Unary	K-tuple	4
Mismatch penalty	1	Joining penalty	30
Gap penalty	1.00	Window size	32
Gap size penalty	0.33		
Cutoff score	5		
Randomization group	0		
Initial scores to save	45	Alignments to save	15
Optimized scores to save	0	Display context	10
		SEARCH STATISTICS	

```

Number of residues:      307267536
Number of sequences searched: 358124
Number of scores above cutoff: 4841

```

Cut-off	raised	to 26
Cut-off	raised	to 30
Cut-off	raised	to 33
Cut-off	raised	to 37
Cut-off	raised	to 39
Cut-off	raised	to 42
Cut-off	raised	to 44
Cut-off	raised	to 47
Cut-off	raised	to 49
Cut-off	raised	to 52
Cut-off	raised	to 55
Cut-off	raised	to 58
Cut-off	raised	to 60
Cut-off	raised	to 63

Cut-off raised to 65.
Cut-off raised to 68.
Cut-off raised to 71.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

2 100% identical sequences to the query sequence were found:

Sequence Name	Description	Length	Score	Init.	Opt.	Sig.	Frame
1. HDMMLCPL	Human c-mpl ligand (ML) mRNA,	1795	1795	1795	130.53	0	
2. HSMMLCPL	Human c-mpl ligand (ML) mRNA,	1795	1795	1795	130.53	0	

The list of other best scores is:

Sequence Name	Description	Length	Score	Init.	Opt.	Sig.	Frame
3. HS011025	*** 96 standard deviations above mean *** Human megakaryocyte growth an	1341	1338	1338	96.68	0	
4. HDMTA	*** 87 standard deviations above mean *** Human gene for thrombopoietin	7666	1216	1441	87.64	0	
5. HDMTHROMA	*** 83 standard deviations above mean *** Human thrombopoietin gene, co	6163	1156	1376	83.19	0	
6. HDMTHROMB	*** 76 standard deviations above mean *** Human thrombopoietin mRNA, co	1062	1062	1062	76.23	0	
7. MUSTHROA	*** 48 standard deviations above mean *** Mus musculus thrombopoietin m	1486	694	1155	48.97	0	
8. HSTNSCN	*** 10 standard deviations above mean *** Human mRNA for tenascin.	7286	180	780	10.89	0	
9. HDMHXB	*** 9 standard deviations above mean *** Homo sapiens hexabrachion MRN	7390	180	781	10.89	0	
10. HSHXB	*** 8 standard deviations above mean *** Homo sapiens hexabrachion MRN	7390	180	781	10.89	0	
11. HSTENAS3	*** 9 standard deviations above mean *** Mouse ZP3 gene, encoding zona	1317	177	576	10.67	0	
12. MVSZP3	*** 9 standard deviations above mean *** Adenovirus type 5 fiber prote	2530	166	704	9.85	0	
13. ADRTIBA	*** 8 standard deviations above mean *** Discopryge omata Ca2+ channel	6981	164	701	9.70	0	
14. DYGCALIB	*** 8 standard deviations above mean *** Murine GLUD mRNA for glutamat	2942	159	738	9.33	0	
15. MMGLUD	*** 8 standard deviations above mean *** G.gallus mRNA for integrin be	3482	154	770	8.96	0	
16. GINTB3	*** 7 standard deviations above mean *** Human mRNA fragment for plate	2129	153	362	8.89	0	
17. HSPIGZBI	*** 7 standard deviations above mean *** R.norvegicus beta-1-adrenergic	1645	152	724	8.82	0	
18. RATBIAR	*** 7 standard deviations above mean *** C.porcillus mRNA for preproac	1320	147	596	8.45	0	
19. CEPACRO	*** 7 standard deviations above mean *** X.laeviis POMC-B gene for proo	7770	144	697	8.22	0	
20. XIPOMCB	*** 7 standard deviations above mean *** Human cytochrome P45011E1 (et	14776	144	776	8.22	0	
21. HSCYIIE	*** 7 standard deviations above mean *** C.eellipoida ribcl, rps14, tr	9705	143	735	8.15	0	
22. CHIC27	*** 7 standard deviations above mean *** C.eugametes l1637 gene.	2935	142	765	8.08	0	
23. CELIC37	*** 7 standard deviations above mean *** H.sapiens CAT mRNA.	2461	141	758	8.00	0	
24. HSCATMR	*** 7 standard deviations above mean *** Bacteriophage S13 circular DN	5386	141	486	8.00	0	
25. S13CG	*** 7 standard deviations above mean *** Human endochelial nitric-oxid	23142	141	781	8.00	0	
26. HSGNOS48	*** 7 standard deviations above mean *** Rattus norvegicus interluekin	2598	140	519	7.93	0	
27. RATIIRRC	*** 7 standard deviations above mean *** Human RHD blood group antigen	1354	139	593	7.85	0	
28. HDMRDANTI	*** 7 standard deviations above mean *** H.sapiens mRNA for rheus pol	1545	139	672	7.85	0	
29. HSRHII	*** 7 standard deviations above mean *** H.sapiens mRNA for rheus pol	2790	139	747	7.85	0	
30. HSRHIIIT	*** 7 standard deviations above mean *** Yeast (S.douglasi) NAM2 gene	4072	139	648	7.85	0	
31. SDNAM2G	*** 7 standard deviations above mean *** Rattus norvegicus phospholipa	7203	139	761	7.85	0	
32. RATPHOSPHB	*** 7 standard deviations above mean *** T.thermophilus gene for lyeyl	1738	137	694	7.70	0	
33. TLYVS	*** 7 standard deviations above mean *** T.thermophilus gene for lyeyl	1738	137	694	7.70	0	

34. CTE64CUD	C.coturnix clusterin gene.	6301	137	753	7.70	0
35. HSC1352	Human proto-oncogene (BCU3) g	4776	135	782	7.56	0
36. MDS19R	Mouse interluekin 9 receptor	3020	134	760	7.48	0
37. HDMENIA	Human mRNA.	3132	134	764	7.48	0
38. CHKA3A1	Gallus domesticus beta A3 cry	5327	134	724	7.48	0
39. HSPAIIV	Human pro-alpha-1 (V) collage	7138	134	781	7.48	0
40. HSPVAF	Human von Willebrand factor p	21033	134	756	7.48	0
41. ADRCOMPEN	Maedadenovirus h5 gene, compl	35935	134	776	7.48	0
42. HDMCYTNEWB	Homo sapiens (clone pat 744)	640	133	247	7.41	0
43. HSCYTNEWB	Homo sapiens (clone pat 744)	640	133	247	7.41	0
44. HDMPIIB	Human platelet glycoprotein I	703	133	304	7.41	0
45. HSGPIIB	Human platelet glycoprotein I	703	133	304	7.41	0

1. US-08-223-263-2 (1-1795)

HDMMLCPL Human c-mpl ligand (ML) mRNA, complete cds.

LOCUS HDMMLCPL 1795 bp se-mRNA PRI 07-JAN-1995
DEFINITION Human c-mpl ligand (ML) mRNA, complete cds.
ACCESSION L33410
KEYWORDS c-mpl ligand.
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens

REFERENCE
AUTHORS
de Sauvage, F.J., Haas, P.E., Spencer, S.D., Malloy, B.E., Gurney, A.L.,
Spencer, S.A., Darbonne, W.C., Henzel, W.J., Wong, S.C., Kuang, W.-O.,
Oles, K.J., Hultgren, B., Solberg, L.A.Jr., Goeddel, D.V. and
Eaton, D.L.
1 (bases 1 to 1795)
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
Eutheria; Primates; Haplorhini; Catarrhini; Homidae.

TITLE
JOURNAL
MEDLINE
COMMENT
FEATURES
source
stimulation of megakaryocytopoiesis and thrombopoiesis by the c-mpl
ligand [see comments]
Nature 365 (6481), 533-538 (1994)
94261202
NCBI gi: 506826
location/Qualifiers
1..1795
/organism="Homo sapiens"
/sequenced_mol="cDNA to mRNA"
216..278
/gene="ML"
/codon_start=1
216..1277
/gene="ML"
/note="erythropoietin homology domain bp 66..522; NCBI
gi: 506827"

CDS
sig_peptide
sig_peptide
/product="c-mpl ligand"
/translation="MELTELLVLLVLLTARLTSSAPACDLRVSLRDSHVIH
SRISOCEVHPIPPVAVDFSIEMKTEKFAODIIGAVTLLEGMAKRGD
GPTLSSILGQSGQVRLILGALISLGTQVPOGRTAKHPNFIISPHILRGV
RFLNIVGSLTICVRAPPTAVSRISLVITLNEIPNRTSGLETFVARSATISGL
IKWQGRAPKIPGLINDTSIDQIPGYINRIHELINGRGLFPPSRRTGADISS
TSPDLSIPNIDGYSPSPTEHPTEGYTLFPLPTLPVVOVLPALDPDSAPTPP
TSPLNTSYHSQNLISQEG"

mat_peptide

BASE COUNT 420 a 586 c 382 g 407 t

ID	Accession	Species	Gene	Length			
1370	1380	1390	1400	1410	1420	1430	1440
AC	CTTCTCTGAA	ACCCAAAGCCCTCGTAA	AAAGGATACACAGAC	CTGAAAGGAAATATTTTCACTGTA			
XX	11111111						
XX	ACTTCTCTGAA	ACCCAAAGCCCTCGTAA	AAAGGATACACAGAC	CTGAAAGGAAATATTTTCACTGTA			
1370	1380	1390	1400	1410	1420	1430	1440
CA	TATTAACCTTCAGAA	ACTTATTTTAAAGCTATCAGCA	ATATCTATCATGACGACGCTA	AGCTCTTTGCTCT			
XX	11111111						
XX	1450	1460	1470	1480	1490	1500	1510
AT	TTTCTGCAAGAA	ATTGCACTACTACTATCTCTAC	ACTGCTTTTCTGTATAC	CTGCAAGGCTG			
XX	11111111						
XX	ATTCTGCAAGAA	ATTGCACTACTACTATCTCTAC	ACTGCTTTTCTGTATAC	CTGCAAGGCTG			
1520	1530	1540	1550	1560	1570	1580	
GG	CTGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
XX	11111111						
XX	GGCTGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
1590	1600	1610	1620	1630	1640	1650	
GG	GTGGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
XX	11111111						
XX	GGGTGGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
1590	1600	1610	1620	1630	1640	1650	
GG	GTGGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
XX	11111111						
XX	GGGTGGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
1590	1600	1610	1620	1630	1640	1650	
GG	GTGGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
XX	11111111						
XX	GGGTGGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
1590	1600	1610	1620	1630	1640	1650	
GG	GTGGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
XX	11111111						
XX	GGGTGGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
1590	1600	1610	1620	1630	1640	1650	
GG	GTGGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
XX	11111111						
XX	GGGTGGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
1590	1600	1610	1620	1630	1640	1650	
GG	GTGGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
XX	11111111						
XX	GGGTGGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
1590	1600	1610	1620	1630	1640	1650	
GG	GTGGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
XX	11111111						
XX	GGGTGGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
1590	1600	1610	1620	1630	1640	1650	
GG	GTGGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
XX	11111111						
XX	GGGTGGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
1590	1600	1610	1620	1630	1640	1650	
GG	GTGGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
XX	11111111						
XX	GGGTGGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
1590	1600	1610	1620	1630	1640	1650	
GG	GTGGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
XX	11111111						
XX	GGGTGGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCCCTTT			
1590	1600	1610	1620	1630	1640	1650	
GG	GTGGCGCTCGCAG	AGTTGAACAAGGAGACGAC	TACCTTGACGTAAGAAACAG	AGAAAGGTAATTTCC			

[illegible]

[illegible][illegible]


```
|||||
AGAGCAAAATTCCTGCTGCTGTAACCAACCTCCAGCTCCCTGAGCAAAATCCCGGATACCTTAACAGG
120 730 740 750 760 770 780
970 980 990 1000 1010 1020 1030
ATACAGAACTCTTAATGAGAACTGCTGAGCTCTTCTGAGACCTCAGCAGAGACCTTAGAGACCCGAGC
|||||
ATACAGAACTCTTAATGAGAACTGCTGAGCTCTTCTGAGACCTCAGCAGAGACCTTAGAGACCCGAGC
790 800 810 820 830 840 850
1040 1050 1060 1070 1080 1090 1100
ATTCTCTGAGAACTCAGACAGACAGCTCCGACCACTCCAGCTGATATTCCTTCCGCAACC
|||||
ATTCTCTGAGAACTCAGACAGACAGCTCCGACCACTCCAGCTGATATTCCTTCCGCAACC
860 870 880 890 900 910 920
1110 1120 1130 1140 1150 1160 1170
CATCTCTCTGAGAACTCAGACAGCTCCGACCACTCCAGCTGATATTCCTTCCGCAACC
930 940 950 960 970 980 990
1180 1190 1200 1210 1220 1230 1240 1250
CCCCCTCTCTGAGAACTCAGACAGCTCCGACCACTCCAGCTGATATTCCTTCCGCAACC
|||||
CCCCCTCTCTGAGAACTCAGACAGCTCCGACCACTCCAGCTGATATTCCTTCCGCAACC
1000 1010 1020 1030 1040 1050 1060 1070
1260 1270 1280 1290 1300 1310 1320 1330
TCCCAAAATCTCTGAGAACTCAGACAGCTCCGACCACTCCAGCTGATATTCCTTCCGCAACC
|||||
TCCCAAAATCTCTGAGAACTCAGACAGCTCCGACCACTCCAGCTGATATTCCTTCCGCAACC
1080 1090 1100 1110 1120 1130 1140
1330 1340 1350 1360 1370 1380 1390
CCTTCCCTGAGAGGCGCCCTGAGAGACAACTGACAAATTCCTTCTCTGAGAACTCCGAGCCTG
|||||
CCTTCCCTGAGAGGCGCCCTGAGAGACAACTGACAAATTCCTTCTCTGAGAACTCCGAGCCTG
1150 1160 1170 1180 1190 1200 1210
1400 1410 1420 1430 1440 1450 1460
GTAAGAGGATACAGAGAACTGAGAACTGAGAACTGAGAACTGAGAACTGAGAACTGAGAACTGAG
|||||
GTAAGAGGATACAGAGAACTGAGAACTGAGAACTGAGAACTGAGAACTGAGAACTGAGAACTGAG
1220 1230 1240 1250 1260 1270 1280
1470 1480 1490 1500 1510 1520 1530
TTTAACTATACAGAACTGAGAACTGAGAACTGAGAACTGAGAACTGAGAACTGAGAACTGAGAACT
|||||
TTTAACTATACAGAACTGAGAACTGAGAACTGAGAACTGAGAACTGAGAACTGAGAACTGAGAACT
1290 1300 1310 1320 1330 1340
4. US-08-223-263-2 (1-1795)
HDMTA Human gene for thrombopoietin.
LOCUS HDMTA 7666 bp ds-DNA PRI 23-NOV-1994
DEFINITION Human gene for thrombopoietin.
ACCESSION D32046
KEYWORDS thrombopoietin.
```

```
SOURCE Homo sapiens blood DNA, clone lambdaHGT1.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
AUTHORS Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
1 (bases 1 to 7666)
TITLE Sohma, Y., Akahori, H., Seki, N., Horii, T.-A., Ogami, K., Kawamura, K.
and Miyazaki, H.
JOURNAL Molecular cloning and chromosomal localization of the human
MEDLINE thrombopoietin gene
COMMENT FEBS Lett. 353, 57-61 (1994)
95010765
Submitted (08-Jul-1994) to DDBJ by:
Yoshiaki Sohma
Pharmaceutical Research Laboratory
Kirin Brewery Co., Ltd.
2-2 Souja-machi 1 chome
Maebashi
Gunma 371
Japan
Phone: 0272-54-8618
Fax: 0272-52-2307.
NCBI gi: 577319
FEATURES
source location/Qualifiers
1..7666
/organism="Homo sapiens"
/sequenced_mol="DNA"
/tissue_type="blood"
<1420..1486
/number=1
/label=exon1
3157..3314
/number=2
/label=exon2
join(3302..3314, 3546..3673, 3960..4046, 5979..6146,
6383..7048)
/note="NCBI gi: 577320"
/codon_start=1
/product="thrombopoietin"
/translaction="MELTLLVVMILLTARLTSSPAPACDLRVLSKLRSDSHVH
SRLSQCEVHRLPTVLLPAVDLSLEMTQRETRADIIAGAVLLLEGMAAGOL
GPTCLSLILGOLGOVALLGALISLIGTOLPPOGRTAHKDPNARFISFQHLIRGV
RFLMLVGGSTLCVRAPPTAVSRISLVLTINELPNRTSGLETFNFAARTGSGT
LKMOGGRARIPGILNOTSPSLDIDPGYINRIHEILNTRGLFPGPSRRITGAPDIS
GTSDTGSLPNIDPGISPSPTPTGQYTLFPLPTLPVQLHPILDPSTPTPT
TSPLINTSYTHSONLSOEG"
3546..3673
/number=3
/label=exon3
3960..4046
/number=4
/label=exon4
5979..6146
/number=5
/label=exons
6383..7576
/number=6
/label=exon6
BASE COUNT 1897 a 2154 c 1807 g 1808 t
ORIGIN
```


[illegible]

FEATURES	NCBI gi: 533214	Location/Qualifiers
source	1..6163	/organism="Homo sapiens" /sequenced_mol="DNA" <1..158
exon		/number=1
intron		159..1826 /number=1
exon		1827..1984 /number=2
CDS		join(1972..1984,2216..2343,2630..2716,4649..4816, 5053..5718) /note="NCBI gi: 533215" /codon_start=1 /product="thrombopoietin" /translation="MELTELLLVMLLTARLTLSPPADCLRYLSKLRDLSVLT SRISQCEVHPDPTVLLPADVFSLGSKWOMETKADILGAVTLLEGVAAAGOL GPRCLSLGOLSGOVRLLIALOSLGTOIPGGRTAKDPNAFLISFOILRGKVG RLFMIGQSGTICLVRAPPTTAVPSRTSIVLTINELNRTSGILETFTNLSARTGSGI LKMQGFRAKIRIGLINQNTSRSLDQIPCTAIRIHEILNTRGLEPFGSRRTICADYISS GRIDTSLSPNLDGISEPSTHPPTGGITFLPPLPTPTVQLHLPDPSAPPTP TSELNTSTYSHQNSLDSRG"
intron		1985..2215 /number=2
exon		2216..2343 /number=3
intron		2344..2629 /number=3
exon		2630..2716 /number=4
intron		2717..4648 /number=4
exon		4649..4816 /number=5
intron		4817..5052 /number=5
exon		5053..>6163 /number=6
BASE COUNT	1524 a 1706 c 1435 g 1498 t	
ORIGIN		
Initial Score =	1156	Optimized Score = 1376
Residue Identity =	81% Matches	Significance = 83.19
Gaps	86	Mismatches = 250
		Conservative Substitutions = 0
	X	10 20 30 40 50 60
	TCCTCTACCATCTGCTCCACAGAGGC-TGCCTGCTGTGACATGGGTCCTGAGCCCTT	
	1 111 111 111 111 111 111 111 111 111 111 111	
	CTGAACAACAGTGTAAAGCAGCGCTG-GACTTGAAGAGACAGGAGATTTTTAAACACAGAGCGTATGACCAG	
	4420 4430 4440 4450 4460 4470 4480	
	CT-----CAACCCGGATAG-ATTCTCTACCCCTTGGCCGG-CCTTGGCCCCACCC--TACTTGGCCAGAG	
	1 11 11 11111 111 11 1111 11 1111 11 11 111 111 11	
	CTCTGGGAGACTGTGAAGAAATTCCT-GCCCTGGGTGGGACCTTGTCTCTCCAGTTCACGCTGTATG	
	4490 4500 4510 4520 4530 4540 4550	
	130 140 150 160 170 180	
	-TGCAGAGCGCTTAAGCGCGCTTC-ATGGC-CCACGAGAGGATTTCAGGGGAGAGGCCCC--CAAAACAGGAGC	

[illegible]

6. US-08-223-263-2 (1-1795)
HUMTHROMB Human thrombopoietin mRNA, complete cds.

LOCUS	DEFINITION	ACCESSION	KEYWORDS	SOURCE	ORGANISM
HDMTHROB	Human thrombopoietin mRNA, complete cds.	U36052	thrombopoietin, Homo sapiens CDNA to mRNA.		
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homnidae.					
REFERENCE					
AUTHORS	Foster,D.C., Sprecher,C.A., Grant,F.J., Kramer,J.M., Kilper,J.I., Holly,R.D., Whitmore,T.E., Heipel,M.D., Bell,L.A.N., Ching,A.F., McGrane,V., Hart,C., O'Hara,P.J., and Lok,S.				
TITLE	Human Thrombopoietin: gene structure, cDNA sequence, expression, and chromosomal location				
JOURNAL COMMENT	Proc. Natl. Acad. Sci. U.S.A. 91, 13023-13027 (1994)				
FEATURES	NCBI gi: 533216				
source	Location/Qualifiers 1..1062				

CDS

```

/note="putative; NCBI gi: 533217"
/codon_start=1
/product="thrombopoietin"
/translacion="MELTELIELVALLILYTLATLTLSPPAPACDLRVLSKLLRDSHVLAH
SRISQCEPVHPLPTPLVPLPAVDFSLGEKQIMETEAQDILGAVLLLEGMARQGLD
GPTCCISLLIGQLSGCYRLLIALDLSLGTQPLPGSGRTAHKQPNALFSPQHLRGV
RFLMLVGGSTLTCVRRAAPPTLVAPSSTLSVLINLDELNRRITSGLETNFTAKRTGGSL
LKMQQGFRAKIPGLINQTSRLDQIPGLINRIHELNRGKGLPEPSRRTTGAADISSL
GDTDSLTSLPDLQPGVSPSTHPPTGQVTLFPLPTLPVQVLAHLPLDPSATPTPP
TSLPLNTSYTHSONLSQEG"
BASE COUNT      223 a      375 c      236 g      228 t
ORIGIN
Initial Score    = 1062 Optimized Score = 1062 Significance = 76.23
Residue Identity = 100% Matches = 1062 Mismatches = 0
Caps             0 Conservative Substitutions = 0

```

210 X 220 230 240 250 260 270
CCGCGCAGATGGAGCTGACTGAATTGCTCTCGTGTGTCATGCTTCTCCTAACTGCAGGCGTAACGCTTTC

ATGAGCTACTTAATGGCTTCCTCGTGGCATGGCTTCCCTAACTGAAGCGTAAACGGCTGTC
 X 10 20 30 40 50 60
 280 290 300 310 320 330 340
 CAGCCCGGGCTCTCTTCCTTCCTGTATACCTCCGAGATCTCTCAGTAAATGCTTGGTATCTCCATGCCATGTCTTCACAG
 350 360 370 380 390 400 410 420
 CAGACTGACCCAGTCCGACAGAGGCTTACACCTTTCGGCTCAACACTGTCTGTCTGCTCGCTGTGCACTTAACTTACCTTT
 140 150 160 170 180 190 200
 CAGACTGACGACAGTCCGACAGAGGCTTACACCTTTCGGCTCAACACTGTCTGTCTGCTCGCTGTGCACTTAACTTACCTTT

GGGAGATGAAAA 430
GGGAGATGAAAA 440
GGGAGATGAAAA 450
GGGAGATGAAAA 460
GGGAGATGAAAA 470
GGGAGATGAAAA 480
GGGAGATGAAAA 490
GGGAGATGAAAA 500
GGGAGATGAAAA 510
GGGAGATGAAAA 520
GGGAGATGAAAA 530
GGGAGATGAAAA 540
GGGAGATGAAAA 550
GGGAGATGAAAA 560
GGGAGATGAAAA 570
GGGAGATGAAAA 580
GGGAGATGAAAA 590
GGGAGATGAAAA 600
GGGAGATGAAAA 610
GGGAGATGAAAA 620
GGGAGATGAAAA 630
GGGAGATGAAAA 640
GGGAGATGAAAA 650
GGGAGATGAAAA 660
GGGAGATGAAAA 670
GGGAGATGAAAA 680
GGGAGATGAAAA 690
GGGAGATGAAAA 700
GGGAGATGAAAA 710
GGGAGATGAAAA 720
GGGAGATGAAAA 730
GGGAGATGAAAA 740
GGGAGATGAAAA 750
GGGAGATGAAAA 760
GGGAGATGAAAA 770
GGGAGATGAAAA 780
GGGAGATGAAAA 790
GGGAGATGAAAA 800
GGGAGATGAAAA 810
GGGAGATGAAAA 820
GGGAGATGAAAA 830
GGGAGATGAAAA 840
GGGAGATGAAAA 850
GGGAGATGAAAA 860
GGGAGATGAAAA 870
GGGAGATGAAAA 880
GGGAGATGAAAA 890
GGGAGATGAAAA 900
GGGAGATGAAAA 910
GGGAGATGAAAA 920
GGGAGATGAAAA 930
GGGAGATGAAAA 940
GGGAGATGAAAA 950
GGGAGATGAAAA 960
GGGAGATGAAAA 970
GGGAGATGAAAA 980
GGGAGATGAAAA 990
GGGAGATGAAAA 1000

GGGATGTAGTCACACGGGGACACTGGGACCATTGCCCTCCTCCATCCCTCGGGGACGTTTGTGC
500 510 520 530 540 550 560

GGGATGTATGGCACACGGGGACACTGGGACCATTGCCCTCCTCCATCCCTCGGGGACGTTTGTGC
570 580 590 600 610 620 630

GGGATGTATGGCACACGGGGACACTGGGACCATTGCCCTCCTCCATCCCTCGGGGACGTTTGTGC
640 650 660 670 680 690 700

280 290 300 310 320 330 340 350

570 580 590 600 610 620 630
GGGCGTCTCCTCTTGGGGCCGGAGAGCCCTCTGGAAACCAAGTCTCCAGGGCAGACACAG
GGTCCGTCTCTTGGGGCCTGAGAGCCTCTTGGAAACCAAGTCTCCAGGGCAGACACAGC
360 370 380 390 400 410 420

640 650 660 670 680 690 700
TCACAGATCCCATTCGATCTTCTGAGCTTCACACCTGCTCCGAGGAAAGGTGCCTTTCGATGCT
TCACAGATCCCATTCGATCTTCTGAGCTTCACACCTGCTCCGAGGAAAGGTGCCTTTCGATGCT
TCACAGATCCCATTCGATCTTCTGAGCTTCACACCTGCTCCGAGGAAAGGTGCCTTTCGATGCT
430 440 450 460 470 480 490

710 720 730 740 750 760 770 780
TGTAGAGGCTCACCCCTCTCCGACAGCGGCCCCACCAACAGCTGTCCACACACAACCTCTACG
|||||
TGTAGAGGCTCACCCCTCTCCGACAGCGGCCCCACCAACAGCTGTCCACACACAACCTCTACG
500 510 520 530 540 550 560

CCTCACCTGAAGAGAGTCCCAACAGAGCTTGTGATTTGGAGCAAACTTCACTGGCTCAGCCAGAAAC
 790 800 810 820 830 840 850
 CCTCACCTGAAGAGAGTCCCAACAGAGCTTGTGATTTGGAGCAAACTTCACTGGCTCAGCCAGAAAC
 570 580 590 600 610 620 630

860 870 880 890 900 910 920
TACTGGCTCTGGGCTTCTGAACTGGACGACGAGGATTTCAAGCCAGATTCCTGTCTGTCTGAACCAACTT
TTCTGGCTCTGGGCTTCTGAACTGGACGACGAGGATTTCAAGCCAGATTCCTGTCTGTCTGAACCAACTT
TCTGGCTCTGGGCTTCTGAACTGGACGACGAGGATTTCAAGCCAGATTCCTGTCTGTCTGAACCAACTT
640 650 660 670 680 690 700 710

930 940 950 960 970 980 990
CAGTCCCTGGACCAATCCCGGATACCTGACAGGATACAGCAACTTTGATGGAAGCTGGAAGCTTT

530	540	550	560	570	580	590	ACCACTTGGCTGTCAATCCCTCTGGGGAGCTTTTGTGACAGAGTCCGTCTCTCTTGGGGCCCTGACAGAG
490	500	510	520	530	540	550	ACCTCTCGCCTCTCAATCCCTCTGGGAGACGGCTTTCTGGGAGAGTGGCCCTCTTGGGGGGCCCTGACAGAG
600	610	620	630	640	650	660	CCTCTCTGGAAACCGAGCTTCTCTCAACAGGGACAGACACAGCTCAACAAAGATCCCAATGCATCTTCTCTGAG
560	570	580	590	600	610	620	CCCTCTAGAACCCAGCGTTCTCTTACACGGGACAGAACACAGCTCAACAGACCCCATGGCCCTCTTCTTGAAG
670	680	690	700	710	720	730	CTTCCACACACTGTCTCCGAGGAAAGGTGCGTTCTCTATGCTTTGTAGAGGGGTCAACCTCTGCTACCGGCG
630	640	650	660	670	680	690	CTTGGACACACTGTCTTGGGGGAAAGGTGCGCTTCCCTGCTTCTGTGTAGAAAGGTCCACACCTCTGTCTGACAC
740	750	760	770	780	790	800	GGCCCCACCCACACAGCTGTCCCCAGACACAACTCTCTAGTCTCACACTGAACAGAGCTCCCAACAGAG
700	710	720	730	740	750	760	GACCTCCACACACACAGCTGTCCACAGAGAGTACTTCTCACTCTCTCACTCAACAAAGTGTCCCAACAGAG
820	830	840	850	860	870	880	TTCTGGATTGTGGAGACAAATTCACTGTGCTTCAGCCAACTACTGCTCTTGGGTTTCTTGAAGTGGACGA
770	780	790	800	810	820	830	TTCTGGATTGTGGAGAGAACTTCACTGACAGACAACTGCTGGCCCTGACTTTCAGACAGCTTTCAG
890	900	910	920	930	940	950	GGGATTTCAGAGCAAGAT---TCTGCTGTCTGAGACCAAACTTCAAGTCCCTGAGACCAAAATCCCGGATAT
850	860	870	880	890	900	910	GGGATTTCAGAGCAAGATTTACTCTGCTGTCTGATTAATCAACTCCAGTCCCAAGTCCCAATCTGTGATAT
960	970	980	990	1000	1010	1020	CCTGAACAGATATACAGAACTTTGATGATGACTCTGTGACTTTTCTGTGACCCTACGACAGACCCTAAGG
920	930	940	950	960	970	980	CCTGAACAGACACAGGAACTGTGATGATGAAATCAATGAGCGCTTCTGTGAACTCACTTCACTTCACTGAG
1030	1040	1050	1060	1070	1080	1090	AGCCCCGACATTTCTCTAGAGAACTACAGACAGGCTCCCTCCACCAACCTCAAGCTGGATATTTCTCC
990	1000	1010	1020	1030	1040	1050	AGCTTCAGACATCTCGCCCGGAGCTTTCACAAAGGCTCCCTGGGATTCAACTCTCCAGGCTGGAGCTTCTCC
1100	1110	1120	1130	1140	1150	1160	TTTCCCAACCATCTCTCTACTGTGACAGATATAGAGCTTCTCCCTTTCACACACACTGCGACAC--CGTGTGGAG
1060	1070	1080	1090	1100	1110	1120	TTTCTCAACAGCTGTCTCTCTATGAGC---ACAGACCTTCCCTCTTCACTGAGCTTGGCCACACACCAATGAG
1170	1180	1190	1200	1210	1220	1230	-TC-----CAGTTCACACCCCGTCTCTCTGTGACCTTCTGTGTCACAGCCCAACCCCTACACACCTCTTCT
1130	1140	1150	1160	1170	1180	1190	ATCTCACCCCAAGCTTCACACCCCGCTGTTTCTGTGACCTTTCACACACAGTCTTAATCTTAACGGCCCTAATCC

LOCUS	DEFINITION	ACCESSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	TITLE	JOURNAL	MEDLINE	COMMENT
1240	1250	1260	1270	1280	1290					
AAACATCTCCACACCCACCTCCCA-GAATCTCTCTCAGAAAGGTA----	AGTTCTCAGACACTG-CCGAC									
AGTCACAAATGATCCCTCA-TCGCCAGAAATTTCTCAGAAACATGCGCGGCACTGCCCAATGACACCTG										
1200	1210	1220	1230	1240	1250	1260				
ATAGACATGCTCTCANTGACAGCTCCCTCCCTCAGAGCGGCCCTGGAGACAACTGGA-----CAAGAT										
TGAGG-CTTCTCTCGGGAGCA---AGCTTCCC-CAGGAAG--GCTGAGGCGAGCTGCATCTGCTCCAGAT										
1270	1280	1290	1300	1310	1320	1330				
1370	1380	1390	1400	1410	1420	1430				
TTCCCTACTTTCTCTCGAAMCCCAAGCCCTCTGTAAGAGGATACACAGACTGCAAAAGGAATCATTTTTC										
GTTCTGCTTTCACTTA-----AAGCCCTTG-GGAAAGGATTCACAG-----CA										
1340	1350	1360	1370							
CTGTACATTTATTAACCTTCAGAGCTA-----TTTTTTAGCTATACAGCATATCTCATCAGACAGCGTAC--T										
CTGAGAGTGTAAATTTTAGAGCTATTTTATTTTAACTATCAGCATTCATCATCAGACAGCGTACCGAT										
1380	1390	1400	1410	1420	1430	1440				
1510	1520	1530	1540	1550						
CTTTGGCTATTTTCTCGCAGAAATTTTGCACTCATCTATCTCTAC										
CTTTGGCTATTTTGGTATTAATTTGAATTAACACTA										
1450	1460	1470	1480	X						

8. US-08-223-263-2 (1-1795)
 HSTNSCN Human mRNA for tenascin.
 LOCUS HSTNSCN 7286 bp RNA PRI 01-SEP-1992
 DEFINITION Human mRNA for tenascin.
 ACCESSION X56160
 KEYWORDS extracellular matrix protein; tenascin.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
 Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homidae.
 1 (bases 1 to 7280)
 REFERENCE Direct Submission
 Submitted (10-OCT-1990) to the EMBL/GenBank/DBJ databases. L.
 Zardi, I, ISTITUTO NAZIONALE PER LA RICERCA SUL, CANCRO VIALE BENNETTO
 XV, 10, 16132 GENOVA, ITALY
 2 (bases 1 to 7280)
 Stri, A., Carnemoli, B., Saginati, M., Lepini, A., Casari, G.,
 Barile, F. and Zardi, L.
 Human tenascin: primary structure, pre-mRNA splicing patterns and
 localization of the epitopes recognized by two monoclonal
 antibodies
 Nucleic Acids Res. 19 (3), 525-531 (1991)
 91187670
 This sequence conflicts with M24630 (Gulcher J.R. P.N.A.S.
 86:1588-1592(1989)).

[illegible]

LOCUS	HUMHXB	7390 bp ss-mRNA	PRI	08-NOV-1994
DEFINITION	Homo sapiens hexabrachion mRNA, complete cds.			
ACCESSION	M5618			
KEYWORDS	hexabrachion; tenascin.			

SOURCE Human, cDNA to mRNA.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
Nies, D.E., Hemeaath, T.J., Kim, J.H., Gulcher, J.R. and Stefansson, K.
The complete cDNA sequence of human hexabrachion (Tenascin). A
multidomain protein containing unique epidermal growth factor
repeats
J. Biol. Chem. 266 (5), 2818-2823 (1991)
MEDLINE 91131572
REFERENCE 2 (bases 1 to 7390)
AUTHORS Gulcher, J.R., Nies, D.E., Alexakos, M.J., Ravikant, N.A.,
Sturgill, M.E., Marton, L.S. and Stefansson, K.
Structure of the human hexabrachion (tenascin) gene
Proc. Natl. Acad. Sci. U.S.A. 88 (21), 9438-9442 (1991)
TITLE JOURNAL 92052108
COMMENT NCB1 g1: 184483
FEATURES Location/Qualifiers
1..7390
/organism="Homo sapiens"
/sequenced_mol="cDNA to mRNA"
1..592
/partial
/gene="Hxb"
/number=1
/codon_start=1
136..6747
/gene="Hxb"
/map="9q32-q34"
/note="NCBI g1: 184484"
/codon_start=1
/gb_xref="600-120-073"
/product="hexabrachion"
/translation="MGAMTQLAGVLIATLATEGCVLKVIRHKRSQGNATLPER
NQPVFNVHYNIKLIPVSGQSVLSESAHGKDLAPPESESEFOEHTVDGNQVTFH
RINIPRACGCAAPDVKEILSLRELELVLSIRECTAGAGCICOPATGRLDTRP
CSGRNFSTEGCGVCEGKMGKPGNCSPEPCSGHRCIDGOCICDDGTGEDCSQ
LACPDNDGCKVNGVCICTEGVACANDCSREICPCSEHGTCVDGICVCHGFAG
DDCKRP/LCINNCYRGRVCNECYCDEGTGEDCSLLICPNDPFRGRCINGTCCE
GTGEDCGKPTCPHACTGQCEGCGVCEGAGLDCSEKCPADCHNRGCVDRG
ECDDGTGADGELKCPNGSGHRCVNGCGVCEGAGLDCSEKCPADCHNRGCVDRG
EGKVCCEGKGDGSDMSPNDCHGRCVNGCGVCEGAGLDCSEKCPADCHNRGCVDRG
GLVCGGVCEDGTGPDCAELSCPDNCHGRCVNGCGVCEGAGLDCSEKCPADCHNRGCVDRG
CHGGRVCVDCGICGHEGTCIDGCHSCPDNCHGRCVNGCGVCEGAGLDCSEKCPADCHNRGCVDRG
PPKDLVTEVTEETVNLAMNEMKRVETLYVPTTREGGLEMQFRVPGDQSTIIQEL
EPGEYFIRVAILLENKRSIPVSARVATYIPAEGLKRSIKETSEVEEMDPLIAE
TWEIIFRNKNEDEGLITKSLRREPSTYRGTGLAPGOEYELSHIVKNTNPGIKRV
TTTSLDAPSOLEVDVDTITLITWFKPLAELIDGILTYGIDKVPDRITDITFEDN
OYSTGNKPDTEYVSLISRGDMSNPAAEFPTGLDAPRLARVSONDSITLERN
NGKAAIDSYRIKVPISGDPHAEVDVPSQAATTGTTGIRPTEGICVSAKZEK
ESNPATINATELDTPKDLVSEYETSLTLAKTTLAFDRRLNYSILPTGOMGVO
LPRNTSVILGLEPGEYVLLTAEGKRSKSPAHVASTEOAPLENLIVTEVGM
GLRLMTADQAYEHFIIIOEANKVEARNLIVPSLAVDIPGKAATPVYSIG
VIOGRTPVLSAEASTGETENIGSEVVAEYGMALKINTAPAGAEYFIOVEDAT
VEAONLIVPGILSTDPGILKATHTTIRBGVODFTSTPLVSEVILVEEPMGNL
TTEVSMALRLNWTTPDGTDOFTTIOVGPADVEAHNLIVPSLSRMEIPGACAT
PYTTLHGEVRGSHTRPLAVEVTEDELPQGDILAVEVGMGLRLMTADQAYEHF
IOVEDANKVEAONLIVPGILRAVDIPGILAEATPVYSIGVIRGRTPVLSAEASTA
KEPEIGNLWSDITPESFNLSMAWDGIFETFTIETIDSNRLTEVYNSCAERTAA

ISGLPSTDFIVYLSGLAPSRIRKTIISATATTEALPLLENLITSDINPYGFTVSNAS
ENAPDSFIATVVDGSKLDPOEFTLISGTQRKLEGLITGIGYEWMLMIPRASKNP
LRAEIVTAEPEDVNDLIVDADTPDGFRLSWTDEGVPNNFYIKIRDKKQSEPLEITL
LAPERRDPLGLKEATEYEIEIXGISKGRSQTVAIATTMAGSEKVEYISDITENSA
IVSWRAPTAQVESFRITVYITGGTSPVAVDGTQTRVLKLLIPGEVLSIITAKG
FEESEPVSGFTALDPSGLVYANTDSEALARQPAIATVDSYVLSYTERVEPIT
RTVSGNVEVALDIPATEYTLRIFAEKGPQKSSITAKFTTDDSPRLTAEVOS
ETALLTWPRASVTGVLIVYESVDGTVEKVIIVGDTTSYSLADSPSTHTAKIONL
NGPRLSMIOTITFTTIGLILYEPKQSSOMINGDTTSGIYTYINGCKAOLIVFCDM
TSDGCGIYVLRKNGRENTYONKAYAAFGDREBEFLMHIGLNDLAKTITAGQZEL
RVDLRDGETAFAYDYKF SVGDAKTRIKLVKVEIGTADSSA YHNGRSFTDPKQD
SAITNCALSYKCAFWRNCHRNIMGRYDNNHSGOVNWFHWKCHHSIOFAEKRLRP
SNFENLEGRKKRA"
593..2002
/gene="Hxb"
/number=2
/codon_start=1
2003..2266
/gene="Hxb"
/number=3
/codon_start=1
2267..2382
/gene="Hxb"
/number=4
/codon_start=1
2383..2539
/gene="Hxb"
/number=5
/codon_start=1
2540..2809
/gene="Hxb"
/number=6
/codon_start=1
2810..2995
/gene="Hxb"
/number=7
/codon_start=1
2996..3085
/gene="Hxb"
/number=8
/codon_start=1
3086..3349
/gene="Hxb"
/number=9
/codon_start=1
3350..3622
/gene="Hxb"
/number=10
/codon_start=1
3623..3895
/gene="Hxb"
/number=11
/codon_start=1
3896..4168
/gene="Hxb"
/number=12
/codon_start=1
4169..4441
/gene="Hxb"
/number=13

FT	exon	5923..6055
FT	/gene="Hxb"	
FT	/number=22	
FT	/codon_start=1	
FT	6036..6207	
FT	/gene="Hxb"	
FT	/number=23	
FT	/codon_start=1	
FT	6208..6310	
FT	/gene="Hxb"	
FT	/number=24	
FT	/codon_start=1	
FT	6311..6469	
FT	/gene="Hxb"	
FT	/number=25	
FT	/codon_start=1	
FT	6470..6636	
FT	/gene="Hxb"	
FT	/number=26	
FT	/codon_start=1	
FT	6637..7390	
FT	/partial	
FT	/gene="Hxb"	
FT	/number=27	
FT	/codon_start=1	
XX		
SQ	sequence	7390 BP; 1858 A; 1954 C; 2013 G; 1565 T; 0 other;
Initial Score =	180	Optimized Score = 781
Residue Identity =	49%	Matches = 943
Caps	226	Conservative Substitutions = 744
		Mismatches = 0

[illegible][illegible]

[illegible]

	9980	10000	10020	10040	10060	10080	10100	10120	10140	10160	10180	10200	10220	10240	10260	10280	10300	10320	10340	10360	10380	10400	10420	10440	10460	10480	10500	10520	10540	10560	10580	10600	10620	10640	10660	10680	10700	10720	10740	10760	10780	10800	10820	10840	10860	10880	10900	10920	10940	10960	10980	11000	11020	11040	11060	11080	11100	11120	11140	11160	11180	11200	11220	11240	11260	11280	11300	11320	11340	11360	11380	11400	11420	11440	11460	11480	11500	11520	11540	11560	11580	11600	11620	11640	11660	11680	11700	11720	11740	11760	11780	11800	11820	11840	11860	11880	11900	11920	11940	11960	11980	12000	12020	12040	12060	12080	12100	12120	12140	12160	12180	12200	12220	12240	12260	12280	12300	12320	12340	12360	12380	12400	12420	12440	12460	12480	12500	12520	12540	12560	12580	12600	12620	12640	12660	12680	12700	12720	12740	12760	12780	12800	12820	12840	12860	12880	12900	12920	12940	12960	12980	13000	13020	13040	13060	13080	13100	13120	13140	13160	13180	13200	13220	13240	13260	13280	13300	13320	13340	13360	13380	13400	13420	13440	13460	13480	13500	13520	13540	13560	13580	13600	13620	13640	13660	13680	13700	13720	13740	13760	13780	13800	13820	13840	13860	13880	13900	13920	13940	13960	13980	14000	14020	14040	14060	14080	14100	14120	14140	14160	14180	14200	14220	14240	14260	14280	14300	14320	14340	14360	14380	14400	14420	14440	14460	14480	14500	14520	14540	14560	14580	14600	14620	14640	14660	14680	14700	14720	14740	14760	14780	14800	14820	14840	14860	14880	14900	14920	14940	14960	14980	15000	15020	15040	15060	15080	15100	15120	15140	15160	15180	15200	15220	15240	15260	15280	15300	15320	15340	15360	15380	15400	15420	15440	15460	15480	15500	15520	15540	15560	15580	15600	15620	15640	15660	15680	15700	15720	15740	15760	15780	15800	15820	15840	15860	15880	15900	15920	15940	15960	15980	16000	16020	16040	16060	16080	16100	16120	16140	16160	16180	16200	16220	16240	16260	16280	16300	16320	16340	16360	16380	16400	16420	16440	16460	16480	16500	16520	16540	16560	16580	16600	16620	16640	16660	16680	16700	16720	16740	16760	16780	16800	16820	16840	16860	16880	16900	16920	16940	16960	16980	17000	17020	17040	17060	17080	17100	17120	17140	17160	17180	17200	17220	17240	17260	17280	17300	17320	17340	17360	17380	17400	17420	17440	17460	17480	17500	17520	17540	17560	17580	17600	17620	17640	17660	17680	17700	17720	17740	17760	17780	17800	17820	17840	17860	17880	17900	17920	17940	17960	17980	18000	18020	18040	18060	18080	18100	18120	18
--	------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	----

940	950	960	970	980	990	1000
GACCAATCCCGAATACCTAACAAGATACAG--ACTCTGAATGAGACTCGTGA--CTC--TTTCCGGA						
GA--CACTCTGTGA--AGTGAATGAGATCTCTAAGACTTCGCTTTTAATACCTGGAGATATCTTCGGA						
2490	2500	2510	2520	2530	2540	2550
CCCTACAGGAGACCCCT--AG---GAGGCCCGGACATTTCCTTAGAGAAATACAGACAGCTCCCGACACACC						
ATATGATTAAGAAGATGAGGAGAGATACACA--AAAGCTAGAG--AGGCCAGAGACTCTTACGGGCAAC						
2560	2570	2580	2590	2600	2610	2620
1080	1090	1100	1110	1120	1130	1140
CAACTCCACCCGTGAATATCTCTCTCCCAACCAACCTCCCTACTAG--GACAGTATAGCTCTCCCTCTTCT						
TGCTCTAGCTCCGGGCAAGATATGAGATATCTC--TGACATATGAAAAACAATACCGGGGGCCCTGGCC						
2630	2640	2650	2660	2670	2680	2690
1150	1160	1170	1180	1190	1200	1210
CACCAACCTTGCCACACCCCTGTGTGCAGCTCCACACCCCGTCTCTACACCTTCTCTCCACAGCCACCC						
TGAAGAGGTGACACACAC--ACGGT--TGAT--GCCCCAG---CCAGATCAGAGTCAAGATATCAGACAG						
2700	2710	2720	2730	2740	2750	2760
1220	1230	1240	1250	1260	1270	1280
CTACCAACCCCTCTTTTAAACATCTAC--ACCACATCCACAG--ATCTGTCTCAGAAAGGTA---AGCTTC						
--ACCA--CTGCCT--TGATCACCCTGTTCAAGCCCGCTGGGTAGATCATGATGAGCTGACCTACAGCACA						
2770	2780	2790	2800	2810	2820	
1290	1300	1310	1320	1330	1340	
TCAGACAC--TGCC---GAC---ATCACATTTG--TCTCA---TGAC--AGCTCCCTTCCCTGACGGGCGCC						
TCAAAGCTGCCAGGAGACCCCTACACATCATCTACACAGAGACAGAACAGT--ACTCCATGGGAGAC						
2830	2840	2850	2860	2870	2880	2890
1350	1360	1370	1380	1390	1400	
CTGGAGAC---AACT--GGACAAATTTCCATCTTTCTCTGTAACCAACCAAGCCCTGTAA--AAGGATAT						
CT--GAACTCTGACATGATGATCAGAGTCTCCCTC--ATCTCCGCA--GAGGTGACATGTCACAAACCAGCA						
2900	2910	2920	2930	2940	2950	2960
1410	1420	1430	1440	1450	1460	1470
CACAG--GAC--TGA--AAAGGATCATTTTCTACTGTACATTAATTAACCTTCAGAACTATTTTTTA--AGC						
CAAGAGACCTTTCACACAGGGCCTGATGTCGCCAGGA--AT-----CTTC--GAGCTGTTTCCAGCAGCA						
2970	2980	2990	3000	3010	3020	3030
1480	1490	1500	1510	1520	1530	1540
TATACGAATATCTATACAGACAGCTACTCTTTTG--GTCTATTTTC---TGCAGAAATTTGCACATCACTGA						
3040	3050	3060	3070	3080	3090	3100
TACAGCATACCCCTGATGATGAGATGGCAGAGCAGCTATTGACAGTTACAG--AATTAGATGCCCCCA						
1550	1560	1570	1580	1590	1600	1610
TTCTCTACATGCTCTTTTCTG--TGATATACCTTCAAG--GCTGGGCTGGCTGGGAGTTGAACAGAGGAG						
T--CTCTGGGAGCACACGCTGATGGTATGTTCAAGAGCC--AACAGCG--ACA---ACCAAAACCC						
3110	3120	3130	3140	3150	3160	

[illegible]

12. US-08-223-263-2 (1-1795)

Mouse ZP3 gene, encoding zona pellucida sperm-bind

LOCUS	MUS2P3	1317 bp ss-mRNA	ROD	15-DEC-1988
-------	--------	-----------------	-----	-------------

complete cds.

KEYWORDS sperm-binding protein.

pZP3.2, and DNA, clone MZG I.
ORGANISM *Mus musculus*

Eutheria: Rodentia: Myomorpha: Muridae: Murinae
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;

AUTHORS Ringuette, M.J., Chamberlin, M.E., Baur, A.W., Sobieski, D.A. and

TITLE Molecular analysis of cDNA coding for ZP3, a sperm binding protein

JOURNAL Dev. Biol. 127, 287-295 (1988)

COMMENT
Draft entry and computer-readable sequence [1] kindly submitted by
I Doorn 08-MTC-1988

J. Dean, 09-AUG-1988.

NCBI gi: 202462

source 1..1317

mRNA

```
sig_peptide      30.95
/note="7P3 signal peptide (putative)"
```

CDS
/cc00018calc-1
30..1304

/translaln="MassfylLlClllGecgnsolMlLlGgtpvPvssSxv
 AcLaeAlvYvSbLlGtVpCblLlGBeCpRvSvDlVvRfAaLlHecSv
 OmtKdAlvYvSbLlLHDpVgSjLlRtnKvEvpLlEcRvPgNvSjHlOpTmWpR
 AtvSxkEeLlLlHlMennKtSvPflGvAlLlAeTvYvSjLlLofLdHvCA
 PvspLpDnssPflvYdFhGclvDlGsbSfSvAqvPpRvEtLlOpTvYvHfAnSSv
 MtlvYlChkApAPlNolPDKlNACsFNtKsQvMvEdAdlDQcChGnGssSSS

mat_peptide

QFQIHGPRQWSKLVSRNRRAHVIDEADVYGPLILGKANDQVEGWTSAQPTVALGCL
GLATVAFITLAIIVLAVTRKCHSSSYLVSLPQ*

96..1301

```

/note="sperm-binding protein (ZP3)"
/codon_start=1

```

BASE COUNT	273 a	401 c	340 g	303 t
ORIGIN	Chromosome 6.			

Initial Score	=	177	Optimized Score	=	576	Significance	=	10.67
Residue Identity	=	49%	Matches	=	704	Mismatches	=	530
Gaps	=	196	Conservative Substitutions				=	0

180 X 190 200 210 220 230 240
 GCCCAACAGGAGGCACGCCAGCAGCACCCCGCCAGA--ATGGAGTGACTTAATTGCTTCCTGTGG
 ||| ||| | | | | | | | | | |
 CTGAGC----CGAGTGTAC-TCGAGCGGAGCACAATGGC-TCAAGTAAIT--TCCTCTCC
 X 10 20 30 40 50

TCATGCTCTTCCCAACTGCAGAGCCCAAGCGTCCA--GCCGGGCTCCCTGGTGTGACACCGAGCCCT
 250 260 270 280 290 300 310
 TTTCTCTCGCTGCTGAGGCCGCCAGCTGCATTCGACCTGTGGCTTTG---CGGGT---
 60 70 80 90 100 110
 CAGTAACTGCTCTTGAGACTCCCATGTCTTCACAGACACTGAGCCAGTGCACAGAGGTACACCTTTGCC
 320 330 340 350 360 370 380
 ---GGA-ACTCC---CACCCAGATGGGGGCTCATTA-CGTCTAAGGTG---GAG---TGTCTGAAGC-
 120 130 140 150 160 170

[illegible]

460 470 480 490 500 510
CAAGCACAAGCATTTCTGGAGC--ACTGCCT-TCTCTCT--GAG-GAG--TATGG--CAG--CAC
|||||
CTTGCTCAGAAGGTTCACGCCCGGTTCCTGTATCCACACGTGTAGCTTAAGCCCATTTGCAC

GAGGACAACTGG--GCACCACATTG-CCTTCAT-CCCTCCTG---GGCGACTTTTCT-GGACAGATCCGTCT
GAGTGCA G C A G C A G G G T G C A G A T G A C G A A G A T C C C T G S T A C A G C A C A C T C T A C T C A G A C C C T G C G
320 330 340 350 360 370 380

580
CTCCTTGGGAGCCCGACAGACTTCCTTG---AAACCAG-CTTCTCTCACAGGAGCAGACCAAGATCACAAG
+ + + + +
590
+ + + + +
600
+ + + + +
610
+ + + + +
620
+ + + + +
630
+ + + + +
640
+ + + + +
650
CCTGTGAATGCCCTGTGCATCATCTC--AGGACTACACGTTGGAGGTACCACTTC-AATGCC-GATACCCCAAC

[illegible]

TAGAG-----GGTCCACCTTGTCGATGAGCGAGGCCACCACACAGCTGTCCCAGCAAACTCTCT
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
GAGAGAAACTGCCTTTCTCTCTTCGCGCTATGAGAGAACTGGAATA-CTG-----AG-AAATTCGGCTCC
530 540 550 560 570 580

[illegible]

850
CGACACTGCTGCCTTGGGCTTTGAATGTGCAGCAGGANTCA---GAGCCAAAGTTCCTGCTGTCTGA
860
| | | | |
870
| | | | |
880
| | | | |
890
| | | | |
900
| | | | |
910
-CAG-----CTGTTTGTGACACACTG-CGTGGC--CAGCGCTTACACTTTGTC-AGA--CCCCAATCCTCC
660
670
680
690
700
710

ACCAAC-CTCAGGTCCTGGACCAATCCCGGATCC---TGAAAGGATTAAGAATCTT----GAA
920 930 940 950 960 970
| | | | | | | | | | | | | | | | | |
CCCATACTTCA--TC-GTGCAC---TTCCAGCGTCCCTGTGGATGGCTCTAATTGAAGGCTTTGGCAC
720 730 740 750 760 770 780

980
-TGAAC T -GTTGACTCTTTTC TGT -GACCTCAGCAGAGCCCAAGA GCGCCGCAATTTCTCAGAAC
| | | | |
990
TTTGACGTCCCAAGCCCCGGCCAAAGCTCTC --- CAGTTTACGGTGATGTATTTCATTTTTGCCACAGC
790 800 810 820 830 840 850

1050 1060 1070 1080 1090 1100 1110
ATCAGACAGAGCTCCCTGGCACCAACTCAGCGTGATTTCTTCCCAACC-CATCTCTACTG
||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
TCGCAATTAGGC---CTACATC---ACCTGCCATCT-CAAAGTGGCCAGCATTAACAGATCCCCGATTAG
860 870 880 890 900 910

[illegible]

	1200	1210	1220	1230	1240	1250	1260
ACCCCTTCG-C	CTCCAA	GCGCCAC	CCCCCTTC	TAAACAG	ATCTCA	CCACCCCA	--GAAT
ACATCTGATTC	CTCAGC	CAATGCG	-AAGTAG	TAAATCA	GCCTTCA	CAGTCCAG	TTCAGTAGGACCC
990	1000	1010	1020	1030	1040	1050	

[illegible]

1350 1340 1350 1360 1370 1380
 TGC-----CTTCCTGGAGGGGCCCCCTGGAGAGAACTGAGCAAGTTTCTACTTCTCTCTGAAACCCAAA
 1111 1111 1111 1111 1111 1111
 CCTGATATTCCTTGGAGAGGCCAAGCAGAC--TGTGGA--AGCGTGAGTCG--TTCTGCTAAACCC---
 1130 1140 1150 1160 1170 1180

1390	1400	1410	1420	1430	1440	1450
GCCTTGTTAAAGGATACACAGACTGAAAAGGCAATATTTTCACTGT-AAG-TTATAAAGCTTCAGAA						
-CTGTGGCTCTTGGGTT-----AGGCTGGCCAGACAGTGCATTCCTGACCCGCGACGATATAGCTT-----						
1190	1200	1210	1220	1230	1240	1250
1460	1470	1480	1490	1500	1510	1520
GCATTTTTTAAAGCATACAGCAATCTCATTCAGACAGCTAGCTCTTGGCTATTTTCTGCAGAAATTTGG						
GC---TGTACACAGGAAGTGCATCTCTCTTACACTTGTATCCCTTCG-CAG-TTAAAGAGAAACTCA						
1260	1270	1280	1290	1300	1310	X
1540	1550					
CAACTCACTG						

13. US-08-223-263-2 (1-1795)
ADRFIBA Adenovirus type 5 fiber protein gene, complete cds

LOCUS	ADRF1BA	2530 bp ds-DNA	VRL	15-MAR-1989
DEFINITION	Adenovirus type 5 fiber protein gene, complete cds.			

KEYWORDS	capsid protein; fiber protein.
SOURCE	Adenovirus type 5 DNA, clone pECor1 B
ORGANISM	Mastadenovirus h5

REFERENCE 1 (bases 1 to 2530)

TITLE The sequence of adenovirus fiber: Similarities and differences between serotypes 2 and 5
JOURNAL Virology 161, 549-554 (1987)

MEDLINE 88072096

FEATURES

```

/organism="Mastadenovirus h5"
476..2221
CDS

```

BASE COUNT 762 a 671 c 445 g 652 t
 Origin 3 bp upstream of HpaI site.

Initial Score =	166	Optimized Score =	704	Significance =	9.88
Residue Identity =	49%	Matches =	874	Mismatches =	644
Gaps =	247	Conservative Substitutions			0

150 X 160 170 180 190 200 210
CATGCCCAAGAAATTCAGGGGAGAGCCCCCAACAGGAGCCACGCCAG-CCAGACACCCCGCCAGAT

GTTAACTTGCAACCACTC-----CAAA-AGGG-GTATCTTTTGTCTGTAAAGCAGGCCA-A
 X 10 20 30 40 50

ATGAGGCT--GACTG-AATTGCTCCTCGTGTCAATGTTTCCTACACGCAG--GGTAAC---GGTGTCAC
 |||||
 ATGTCCTACGACAGTAATACACC-----GGACA---CCGC-CTTAAGCTACAGTTGCCAACACAGC-GTCAG
 |||||

280 290 300 310 320 330 340
GCCGCGCTCCCTCCTCGCT-TGTGA-CCTCCAGTCTCACTAAATGC-TTGG--TAGCTCCACTGTCTTCAC
| | | | |
120 130 140 150 160 170 180
AAATTGTGTCATGCTGGGAGAAAAGCCATTAC-CATAAATCGACACTGCTAGAACAACGAAG-GCTGCC

350 360 370 380 390 400
 --CAGCAGCTACGACAGGCCAAGG-----TTACACCT--TGCCTCAACCTTCCTG--CT--GCCGTG
 TTCA-CTACCTCTTAAGACTAGAGATCTGCAACCTTTTAAGC-CTGTGGGCTTCAAAGATCT
 190 200 210 220 230 240 250

910
 410
 420
 430
 440
 450
 460
 470
 480
 490
 500
 510
 520
 530
 540
 550
 560
 570
 580
 590
 600
 610
 620
 630
 640
 650
 660
 670
 680
 690
 700
 710
 720
 730
 740
 750
 760
 770
 780
 790
 800
 810
 820
 830
 840
 850
 860
 870
 880
 890
 900
 910
 920
 930
 940
 950
 960
 970
 980
 990
 1000
 1010
 1020
 1030
 1040
 1050
 1060
 1070
 1080
 1090
 1100
 1110
 1120
 1130
 1140
 1150
 1160
 1170
 1180
 1190
 1200
 1210
 1220
 1230
 1240
 1250
 1260
 1270
 1280
 1290
 1300
 1310
 1320
 1330
 1340
 1350
 1360
 1370
 1380
 1390
 1400
 1410
 1420
 1430
 1440
 1450
 1460
 1470
 1480
 1490
 1500
 1510
 1520
 1530
 1540
 1550
 1560
 1570
 1580
 1590
 1600
 1610
 1620
 1630
 1640
 1650
 1660
 1670
 1680
 1690
 1700
 1710
 1720
 1730
 1740
 1750
 1760
 1770
 1780
 1790
 1800
 1810
 1820
 1830
 1840
 1850
 1860
 1870
 1880
 1890
 1900
 1910
 1920
 1930
 1940
 1950
 1960
 1970
 1980
 1990
 2000
 2010
 2020
 2030
 2040
 2050
 2060
 2070
 2080
 2090
 2100
 2110
 2120
 2130
 2140
 2150
 2160
 2170
 2180
 2190
 2200
 2210
 2220
 2230
 2240
 2250
 2260
 2270
 2280
 2290
 2300
 2310
 2320
 2330
 2340
 2350
 2360
 2370
 2380
 2390
 2400
 2410
 2420
 2430
 2440
 2450
 2460
 2470
 2480
 2490
 2500
 2510
 2520
 2530
 2540
 2550
 2560
 2570
 2580
 2590
 2600
 2610
 2620
 2630
 2640
 2650
 2660
 2670
 2680
 2690
 2700
 2710
 2720
 2730
 2740
 2750
 2760
 2770
 2780
 2790
 2800
 2810
 2820
 2830
 2840
 2850
 2860
 2870
 2880
 2890
 2900
 2910
 2920
 2930
 2940
 2950
 2960
 2970
 2980
 2990
 3000
 3010
 3020
 3030
 3040
 3050
 3060
 3070
 3080
 3090
 3100
 3110
 3120
 3130
 3140
 3150
 3160
 3170
 3180
 3190
 3200
 3210
 3220
 3230
 3240
 3250
 3260
 3270
 3280
 3290
 3300
 3310
 3320
 3330
 3340
 3350
 3360
 3370
 3380
 3390
 3400
 3410
 3420
 3430
 3440
 3450
 3460
 3470
 3480
 3490
 3500
 3510
 3520
 3530
 3540
 3550
 3560
 3570
 3580
 3590
 3600
 3610
 3620
 3630
 3640
 3650
 3660
 3670
 3680
 3690
 3700
 3710
 3720
 3730
 3740
 3750
 3760
 3770
 3780
 3790
 3800
 3810
 3820
 3830
 3840
 3850
 3860
 3870
 3880
 3890
 3900
 3910
 3920
 3930
 3940
 3950
 3960
 3970
 3980
 3990
 4000
 4010
 4020
 4030
 4040
 4050
 4060
 4070
 4080
 4090
 4100
 4110
 4120
 4130
 4140
 4150
 4160
 4170
 4180
 4190
 4200
 4210
 4220
 4230
 4240
 4250
 4260
 4270
 4280
 4290
 4300
 4310
 4320
 4330
 4340
 4350
 4360
 4370
 4380
 4390
 4400
 4410
 4420
 4430
 4440
 4450
 4460
 4470
 4480
 4490
 4500
 4510
 4520
 4530
 4540
 4550
 4560
 4570
 4580
 4590
 4600
 4610
 4620
 4630
 4640
 4650
 4660
 4670
 4680
 4690
 4700
 4710
 4720
 4730
 4740
 4750
 4760
 4770
 4780
 4790
 4800
 4810
 4820
 4830
 4840
 4850
 4860
 4870
 4880
 4890
 4900
 4910
 4920
 4930
 4940
 4950
 4960
 4970
 4980
 4990
 5000
 50

[illegible]

CCCTCCTGGGCGACGTTTTCGAGACAGAGCTCCCTCCCTCGTGCGGCGGCAGACGCTCTGTGAACAGG
| | | | | | | | | | | | | | | | | |
ATTCTCTCCAAATCTAATGGA-ATGTCAGTTCTCT--CTTCCGTGCATCC-----GCACCCA-C

620
 TTCTCCACACAGGGAG--GACACACATCTCAACAGGATC--CCATCCCTCTTCTCGATCTTCCACACCCCT
 630
 TTTCTTATCTGTTTGGAGTGA-AGCGCGAA-CACCTCTGAATATCTTCAACCCCGTATC-CATC
 640
 460 470 480 490 500 510 520

CHCCAGGAAA--GGTGGCTTCTGATGCTGTGAGAGAGACACCCCTGCGACGGCGGCCACACCA
 690 700 710 720 730 740 750
 TACACGGAACGGGCTCTCAACATCTGCTTTCTTA--CTCCGCTTTGATC-----CCCA-----
 530 540 550 560 570 580

CCACGCTGTCACGACAGACCTCTCTATGTCATCAGCAAGACGCTCCCAAC---AG-GATCTTC---
 760 770 780 790 800
 ---ATGGCTTTAAAG-AAGTCCCTCGGGGATCTCTTTGGCC-CTATCCAACTTATTTCTCTCCAA
 590 600 610 620 630 640 650

820 830 840 850 860 870
GGATTCTGGAGACAAACT--TCA-CTGCCTCAGCCAGAACTA--CTGGC--TCTGGGCTTCTGAAGTGG

[illegible]

	ACTACGCTAGCAGAGGCCCTCTTT--TTATAACTCAGCCCAACAATTGGATTTAATACAAACAAGGC	1330	1340	1350	1360	1370	1380	1390
	CTCTCGACAGCTCTTTTTCTGTG-TGATTAACCTTCG--AAAGGCGTGG-GGCTGGCGCT-GGCAGTTGAACAAGAAGG	1400	1410	1420	1430	1440	1450	1460
	CTTGACTTG-----TTTACACCTTCACAAATTCCAAAGGCTTAGGTACGTAACTTAAGCA-CTGCCAAGGGGT	1470	1480	1490	1500	1510	1520	
	AGA-GACATPAC-CTTGAAGTCAGAAAACAAGAAAG--GGTAATTTCCCTTGTCTCAAAATTCAGAGCCCTTC-C	1530	1540	1550	1560	1570	1580	1590
	TATGTTTGAAGCTACAGCCA-TAGCCA-TTATATSCAGAGATGGGC-TTGAATTTGGTTCACCTAATGAC	1600	1610	1620	1630	1640	1650	1660
	AAGCGCCCACATCCCT--TTAGTATCATTTCTAGAGGAGCTGTATCCATTCATTCACATCT--TTACT	1670	1680	1690	1700	1710	1720	1730
	CMAACCAAAATTCCTCTAAACAAAAATTTGGCCAT-GGCTCTAGAT--TTGATTC-AAACAAGGCTATGGTT	1740	1750	1760	1770	1780	1790	1800
	CTTGAGAAATGAAATPAGCTTTCTCTCAGAAAAAAAAAAAAAAAAAAAAA X	1810	1820	1830	1840	1850	1860	1870
	CCTAATACTAGAACACTGCCTTACGTTTGACACACAGAGTGCCATTCATCAGTAGAGAAACA	1880	1890	1900	1910	1920	1930	1940
	DYSCCALB Dyscopyge ommata Ca2+ channel alpha 1 subunit gene	1950	1960	1970	1980	1990	2000	
LOCUS	DYSCCALB 6981 bp ds-DNA VRT 22-JUL-1993							
DEFINITION	Dyscopyge ommata Ca2+ channel alpha 1 subunit gene sequence.							
ACCESSION	J12532							
KEYWORDS	calcium channel alpha-1 subunit.							
SOURCE	Dyscopyge ommata (library: lambda ZAP) adult electric lobe, brain DNA.							
ORGANISM	Dyscopyge ommata Eukaryota; Animalia; Chordata; Vertebrata; Chondrichthyes; Elasmobranchii; Fusedalichii; Neoselachii.							
REFERENCE	I (bases 1 to 6981)							
AUTHORS	Horne,W.A., Ellinor,P.T., Imman,I., Zhou,M., Tsien,R.W. and Schwartz,T.L.							
TITLE	Molecular diversity of Ca2+ channel alpha 1 subunits from the marine ray Dyscopyge ommata							
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 90, 3787-3791 (1993)							
MEDLINE	93248175							
COMMENT	NCBI g1: 290386							
FEATURES	Location/Qualifiers							
source	1..6981 /organism="Dyscopyge ommata" /cell_type="neuron" /dev_stage="adult" /sequenced_moi="DNA" /tissue_type="electric lobe, brain" /tissue_lib="lambda ZAP"							
BASE COUNT	1858 a 1670 c 1767 g 1686 t							
ORIGIN								
Initial Score	164 Optimized Score = 701 Significance = 9.70							

Residue Identity = 50% Matches = 847 Mismatches = 640
Gaps = 204 Conservative Substitutions = 0

```
X
10 20 30 40 50
TCCTCCACCCATCTGCTCCCA-GAG-GC-TGCGCTGTGACTTGGGCTCCGAGCC
TCAGACAGCTTGTGATTCAGAT-TTCTCCAGAGGAGCATCCAGC--AGCA-GTGTGAGCGGAGC--
5390 X 5400 5410 5420 5430 5440
60 70 80 90 100 110 120
TTCTCCACCCGAGATTCCTCAACCTTGGCC-GCGCTTGGCCCACTCTGTGCCAG--ACTGC--
TCAGAG--AGCAATATATCTGAGTTCAGCAACCTTT--CCGACATTTCTGATATATTTGGCTCT
5450 5460 5470 5480 5490 5500 5510
-----AGAGCCTTACGCCCTTCATGCGCCAG-GAAGATTCAGGGGAGAGGCCCA--AACAGGAG
130 140 150 160 170 180
CCCCAAGAGCCTCAGAGTTGACCGGGAAGTCTATGCTGCCCTGATATTACATTAACAAACAG
5520 5530 5540 5550 5560 5570 5580
190 200 210 220 230 240 250 260
CCAGCCGACGACACACCCCGGACAGATGAGCTGATGATTTGCTCTGTGCTATGCTTCTCTACT
TCAGAGAGTAAAGTGCAG--CAGCA--GAGAG--CTGAGTGGCTTGTGAGAGCTCAGAAATCTTCTT
5590 5600 5610 5620 5630 5640 5650
270 280 290 300 310 320
GCAAGCTTACGCTGTTCAGCCCGGCTCC--TCCTGCTTGTGACCTCGAGCTCT-CAGTAACTGCTTCTG
TCAGAGAGTAAAGTGCAG--CAGCA--GAGAG--CTGAGTGGCTTGTGAGAGCTCAGAAATCTTCTT
5660 5670 5680 5690 5700 5710 5720
330 340 350 360 370 380 390 400
GACTCCATGCTCTTCAACAGACTGAGCTGAGGCGGAGGTTTACACCTTGTGCTTACACTGCTCTGCTG
TCAGAGAGTAAAGTGCAG--CAGCA--GAGAG--CTGAGTGGCTTGTGAGAGCTCAGAAATCTTCTT
5730 5740 5750 5760 5770 5780 5790
410 420 430 440 450 460 470
CTGCTGTGACTTTAGCTTGGAGATGAAAAACCCAGATGAGAGCA--AGCA--CAGAGATTTCTG
TCAGAGAGTAAAGTGCAG--CAGCA--GAGAG--CTGAGTGGCTTGTGAGAGCTCAGAAATCTTCTT
5800 5810 5820 5830 5840 5850 5860
480 490 500 510 520 530 540
GAGCAGTGAACCTTCTGCTGAGGAGATGAGCAGGAGGAGACACTGGAGACCACTTCTCTCTCA-TCC
TCAGAGAGTAAAGTGCAG--CAGCA--GAGAG--CTGAGTGGCTTGTGAGAGCTCAGAAATCTTCTT
5870 5880 5890 5900 5910 5920 5930
550 560 570 580 590 600 610
CTCTGGGGAGCTTTCTGAGAGGTCCTCTCTCTTGGGGCCCTGAGAGCCCTCTCTGAA--CCG-A
TCAGAGAGTAAAGTGCAG--CAGCA--GAGAG--CTGAGTGGCTTGTGAGAGCTCAGAAATCTTCTT
5930 5940 5950 5960 5970 5980 5990
620 630 640 650 660 670 680
GCTTCTCTCA--CAGGGCA--GAGACAGAGCTCAGAGATTCAGATTCATCTTCTCTGAG-CTTC-C-AA
TCAGAGAGTAAAGTGCAG--CAGCA--GAGAG--CTGAGTGGCTTGTGAGAGCTCAGAAATCTTCTT
```

CCATCTGGAGAGTCCCTGTCACCGATCCAG--ACCAAG--CCCTATG--AAGCTTCTGTCTACGCTTA
5990 6000 6010 6020 6030 6040 6050

```
680 690 700 710 720 730
CACC--TCTCCAGAGGAAAGTCCGTTCTCTATGCTTGTAGAGGCT--CCA--CCCTTGTGCTGAG
TCAGAGAGTAAAGTGCAG--CAGCA--GAGAG--CTGAGTGGCTTGTGAGAGCTCAGAAATCTTCTT
6060 6070 6080 6090 6100 6110
740 750 760 770 780 790 800
CGGGCCCAACC--CACCACAGCTGT--CCAGAG--ACCTCTATGCTCTCACTGAAAGAGCTCC
TCAGAGAGTAAAGTGCAG--CAGCA--GAGAG--CTGAGTGGCTTGTGAGAGCTCAGAAATCTTCTT
6120 6130 6140 6150 6160 6170 6180
810 820 830 840 850 860 870
AAGAGACTTCTGAT--TGTGAGAGCAAACTTCACTGCTCAGC-CAGA-ACATCTGCTGTGAGCTTC
TCAGAGAGTAAAGTGCAG--CAGCA--GAGAG--CTGAGTGGCTTGTGAGAGCTCAGAAATCTTCTT
6190 6200 6210 6220 6230 6240 6250
880 890 900 910 920 930
TGAA-GTGGAGAGAGATTCAG-ACCAGAG-TTCC-TGCTGTCTGAAACCAACTCA--GGTCCCTG
TCAGAGAGTAAAGTGCAG--CAGCA--GAGAG--CTGAGTGGCTTGTGAGAGCTCAGAAATCTTCTT
6260 6270 6280 6290 6300 6310 6320
940 950 960 970 980 990 1000
-GACC-AAATCC--CGATATCTGAAACAGATTCAGAGACTTGTGATGAACTGTGAGACTTCTCTG
TCAGAGAGTAAAGTGCAG--CAGCA--GAGAG--CTGAGTGGCTTGTGAGAGCTCAGAAATCTTCTT
6330 6340 6350 6360 6370 6380 6390
1010 1020 1030 1040 1050 1060 1070
GACCTTACGAGAGACCTTACGAG-GGCCGAGCA-TTCTCTCAGAGAC-ATCAGACAGAGCTTCTCTGAC
TCAGAGAGTAAAGTGCAG--CAGCA--GAGAG--CTGAGTGGCTTGTGAGAGCTCAGAAATCTTCTT
6400 6410 6420 6430 6440 6450
1080 1090 1100 1110 1120 1130
CCAACTTCCAGC-CTGATATTC--TCTTCCCAACCCATCTCTTACGTGAGAGATATACG-CTCTTCC
TCAGAGAGTAAAGTGCAG--CAGCA--GAGAG--CTGAGTGGCTTGTGAGAGCTCAGAAATCTTCTT
6460 6470 6480 6490 6500 6510 6520
1140 1150 1160 1170 1180 1190 1200
TCTTCCACCCAG--CTTGCCACCC--CTGTGTCTCACTTCAACCCCTCTT--CTGACCTTCTG
TCAGAGAGTAAAGTGCAG--CAGCA--GAGAG--CTGAGTGGCTTGTGAGAGCTCAGAAATCTTCTT
6530 6540 6550 6560 6570 6580 6590 6600
1210 1220 1230 1240 1250 1260 1270
TCAGAGAGTAAAGTGCAG--CAGCA--GAGAG--CTGAGTGGCTTGTGAGAGCTCAGAAATCTTCTT
TCAGAGAGTAAAGTGCAG--CAGCA--GAGAG--CTGAGTGGCTTGTGAGAGCTCAGAAATCTTCTT
6610 6620 6630 6640 6650 6660 6670
1280 1290 1300 1310 1320 1330 1340
G--GTAGGTTCTAGACACTGCGCAGATTCAGATTCATGTA-CACCTTCCCTTCTCTGAGGGGCGCC
TCAGAGAGTAAAGTGCAG--CAGCA--GAGAG--CTGAGTGGCTTGTGAGAGCTCAGAAATCTTCTT
```



```

TTTACATGAAGAAC-----TACTGTCTTCCATTTGAGAGTCTATCTGTGAGTAATTGACCATATATTC
2520      2530      2540      2550      2560      2570      2580
1690      1700      1710      1720      1730      1740      1750
CCATCCCCCTTTACTA-TCAATTCTCAGTGGGACT-CTGATCCCATATTTCTTAACAGATCTTTACTCTTGAG-
1760      1770      1780      1790      1800      1810      1820
AAATCGAGTGTCCGAGAAAAAAGACTTGGGCTCCCTGTT--TGAGATC-TGGCAG-GCTTCGAGCAGCAGAC
2590      2600      2610      2620      2630      2640      2650
AAATGAATTAAGCT--TTCTCTCAGAAAAAAGAAAAAAGAAAAA
1760      1770      1780      1790      1800      1810      1820
CAGTGGCCAGGCTCAACCTGTGTGCTCCCTCCAGAGCTTTCATTCGCTCACC
2660      2670      2680      2690      2700

```

maryh@stic

stdin

NeWSprinter20

Tue Apr 11 13:31:59 1995

NeWSprint 2.5 Rev B

Openwin library 3

NeWSprint interpreter 210.0

NeWSprint 2.5